

FIGURE 1

GGGGCTTCGGCGCCAGCGGCCAGCGCTAGTCGGTCTGGTAAGGATTACAAAAGGTGCAGGTATG
AGCAGGTCTGAAGACTAACATTGTGAAGTTGTAACAGAAAACACTGTTAGAAATGTGGTGGT
TTCAGCAAGGCCTCAGTTCTTCTCAGCCCTGTAATTGGACATCTGCTGCTTCATATT
TCATACATTACTGCAGTAACACTCCACCATATAGACCCGGCTTACCTTATATCAGTGACACTGG
TACAGTAGCTCCAGAAAATGCTTATTGGGCAATGCTAAATATTGCGGCAGTTTATGCATTG
CTACCATTATGTTCGTTATAAGCAAGTTCATGCTCTGAGTCCTGAAGAGAACGTTATCATCAA
TTAAACAAGGCTGGCCTGTACTTGAATACTGAGTTGTTAGGACTTCTATTGTGGCAAACCTT
CCAGAAAACAACCCTTTGCTGCACATGTAAGTGGAGCTGTGCTTACCTTGGTATGGCTCAT
TATATATGTTGTTCAAGACCATCCTTCTACCAAATGCAGCCAAAATCCATGGCAAACAAGTC
TTCTGGATCAGACTGTTGTTGGTTATCTGGTGTGGAGTAAGTGCACCTAGCATGCTGACTTGCTC
ATCAGTTTGACAGTGGCAATTGGACTGATTAGAACAGAAACTCCATTGGAACCCCGAGG
ACAAAGGTTATGTGCTTCACATGATCACTACTGCAGCAGAACATGGTCTATGTCATTCTTCTT
GGTTTCTGACTTACATTGCTGATTTCAGAAAATTCTTACGGTGGAAAGCCAATTACA
TGGATTAACCTCTATGACACTGCACCTTGCCTTATTAAACAATGAACGAACACGGCTACTTCCA
GAGATATTGATGAAAGGATAAAATATTCTGTAATGATTGATTCTCAGGGATTGGGAAAGG
TTCACAGAAGTTGCTTATTCTCTGAAATTCAACCACTTAATCAAGGCTGACAGTAACACT
GATGAATGCTGATAATCAGGAAACATGAAAGAACGCCATTGATAGATTATTCTAAAGGATATCAT
CAAGAAGACTATTAAAAACACCTATGCCTATACTTTTATCTCAGAAAATAAGTCAAAAGACT
ATG

FIGURE 2

<subunit 1 of 1, 266 aa, 1 stop

<MW: 29766, pI: 8.39, NX(S/T): 0

MWWFQQGLSFLPSALVIWTSAAFIFSYITAVTLHHIDPALPYISDTGTVAPEKCLFGAMLNIAAV
LCIATIYVRYKQVHALSPEENVIIKLNKAGLVLGILSCLGLSIVANFQKTLFAAHVSGAVLTFG
MGSLYMFVQTILSYQMOPKIHGKQFWIRLLLVIWCGVSALSMTCSSVLHSGNFGTDLEQKLHW
NPEDKGYVLHMITTAAEWSMSFSFFGFFLTYIRDQKISLRVEANLHGLTLYDTAPCPINNERTR
LLSRDI

Important features:

Type II transmembrane domain:

amino acids 13-33

Other Transmembrane domains:

amino acids 54-73, 94-113, 160-180, 122-141

N-myristoylation sites.

amino acids 57-63, 95-101, 99-105, 124-130, 183-189

FIGURE 3

CGGACGCGTGGCGGACGCGTGGGGAGAGCCGCAGTCCCGCTGCAGCACCTGGAGAAGGCAGACC
GTGTGAGGGGGCCTGTGGCCCCAGCGTGCTGTGGCCTGGGGAGTGGAAAGTGGAGGCAGGAGCCTTC
CTTACACTCGCCCATGAGTTCTCATCGACTCCAGCATCATGATTACCTCCAGATACTATTTTG
GATTGGGTGGCTTCTCATGCGCAATTGTTAAAGACTATGAGATACTGTCAGTATGTTGTACAG
GTGATCTCTCCGTGACGTTGCATTTCTGCACCAGTTGAGCTCATCATCTTGAAATCTTAGG
AGTATTGAATAGCAGCTCCGTTATTTCACTGGAAAATGAACCTGTGTGAATTCTGCTGATCCTGG
TTTCATGGTGCCTTTTACATTGGCTATTTATTGTGAGCAATATCCGACTACTGCATAAACACGA
CTGCTTTCTGTCTTATGGCTGACCTTATGTATTCTCTGGAAACTAGGAGATCCCTTCC
CATTCTCAGCCAAAACATGGATCTTATCCATAGAACAGCTCATCAGCCGGTGGTGTGATTGGAG
TGACTCTCATGGCTCTCTTCTGGATTGGTGTCAACTGCCATACACTACATGTCTTACTTC
CTCAGGAATGTGACTGACACGGATTCTAGCCCTGGAACGGCGACTGCTGCAAACCATGGATATGAT
CATAGCAAAAGAAAGGATGGCAATGGCACGGAGAACATGTTCCAGAAGGGGAAGTGCATAACA
AACCATCAGGTTCTGGGAATGATAAAAGTGTACCACTCAGCATCAGGAAGTGAAAATCTTACT
CTTATTCAACAGGAAGTGGATGCTTGGAAAGAATTAAGCAGGCAGCTTCTGGAAACAGCTGATCT
ATATGCTACCAAGGAGAGAATAGAATACTCCAAACCTCAAGGGAAATATTTAATTCTGGTT
ACTTTCTCTATTACTGTGTTGGAAAATTTCATGGTACCATCAATATTGTTTGATCGAGTT
GGGAAAACGGATCCTGTCAAAAGAGGCATTGAGATCACTGTGAATTATCTGGGAATCCAATTGATGT
GAAGTTTGGTCCAACACATTCTTCATTCTGTGGAATAATCATCGTCACATCCATCAGAGGAT
TGCTGATCACTCTTACCAAGTTCTTATGCCATCTAGCAGTAAGCCTCCAATGTCATTGTCTG
CTATTAGCACAGATAATGGCATGTACTTGTCTCTGTGCTGATCCGAATGAGTATGCCCTT
AGAATACCGCACCATAACTGAAGTCCTGGAGAACTGCAGTTCAACTTCTATCACCGTTGGTTG
ATGTGATCTTCTGGTCAGCGCTCTCTAGCATACTCTTCTTATGGTCACAAACAGGCACCA
GAGAAGCAAATGGCACCTTGAACTTAAGCCTACTACAGACTGTTAGAGGCCAGTGGTTCAAATT
GATATAAGAGGGGGAAAATGGAACCAGGGCCTGACATTATAAACAAACAAAATGCTATGGTAGC
ATTTTACCTTCATAGCATACTCCTCCCCGTAGGTGATACTATGACCATGAGTAGCATGCCAG
AACATGAGAGGGAGAACTAACTCAAGACAATACTCAGCAGAGAGCATCCGTGGATATGAGGCTGG
TGTAGAGGCCAGAGGCCAGAAACTAAAGGTAAAAACTACACTGGAACCTGGCAAGACATGT
CTATGGTAGCTGAGCCAACACGTAGGATTCCGTTAAGGTTACATGGAAAAGGTTATAGCTTG
CCTTGAGATTGACTCATTAAACAGAGACTGTAACAAAAAAAAGGGCGGCCG
ACTCTAGAGTCGACCTGCAGAAGCTGGCCGCATGGCCCAACTGTTATTGCAGCTATAATG

FIGURE 4

MSFLIDSSIMITSQILFFGFGWLFFMRQLFKDYEIRQYVWQVIFSVTFAFSCTMFELIIFEILGV
LNSSSRYFHWKMNLCVILLILVFMVPFYIGYFIVSNIRLLHKQRLLFSCLLWLTMYFFWKLGD
FPILSPKHGILSIEQLISRGGVIGVTLMALLSGFGAVNCPTYMSYFLRNVTDILALERRLLQ
TMDMIISKKRMAMARRTMFQKGEVHNKPSGFWMGIKSVTSASSENLTLIQQEVALEELSRQ
LFLETADLYATKERIEYSKTFKGKYFNLGYFFSIYCVWKIFMATINIVFDRVGKTDPVTRGIEI
TVNYLGIQFDVKFWSQHISFILVGIIIVTSIRGLLITLTKFFYAISSSKSSNVIVLLAQIMGY
FVSSVLLIRMSMPLERYRTIITEVLGELQFNFYHRWFDVIFLVSALSSILFLYLAHKQAPEKQMAP

Important features:

Signal peptide:

amino acids 1-23

Potential transmembrane domains:

amino acids 37-55, 81-102, 150-168, 288-311, 338-356, 375-398,
425-444

N-glycosylation sites.

amino acids 67-70, 180-183 and 243-246

Eukaryotic cobalamin-binding proteins

amino acids 151-160

FIGURE 5

AGCAGGGAAATCCGGATGTCTCGTTATGAAGTGGAGCAGTGAGTGTGAGCCTAACATAGTTCC
AGAACTCTCCATCCGGACTAGTTATTGAGCATCTGCCTCATATCACCACTGGCCATCTGAGGT
GTTTCCCTGGCTCTGAAGGGTAGGCACGATGGCCAGGTGCTTCAGCCTGGTGTGCTTCTCA
TCCATCTGGACCACGAGGCTCTGGTCCAAGGCTTTGCGTCAGAAGAGCTTCCATCCAGGT
GTCATGCAGAATTATGGGGATACCCTGTGAGCAAAAAGGCGAACCAAGCAGCTGAATTTCACAG
AAGCTAAGGAGGCCTGTAGGCTGGACTAAGTTGGCCGGCAAGGACCAAGTTGAAACAGCC
TTGAAAGCTAGCTTGAAACTTGCACTATGGCTGGGTTGGAGATGGATTGTCATCTCTAG
GATTAGCCAAACCCAAAGTGTGGAAAAATGGGGTGGGTGTCCTGATTGAAAGGTTCCAGTGA
GCCGACAGTTGCAGCCTATTGTTACAACCTCATCTGATACTGGACTAACCTGTCATTCCAGAA
ATTATCACCAACCAAAAGATCCCATTCAACACTCAAACACTGCAACACAAACAGAATTATTGT
CAGTGACAGTACCTACTCGGTGGCATCCCCTACTCTACAATACCTGCCCTACTACTCCCTC
CTGCTCCAGCTTCACTCTATTCCACGGAGAAAAAAATTGATTGTCACAGAAGTTTATG
GAAACTAGCACCAGTCTACAGAAACTGAACCAATTGTTGAAAATAAGCAGCATCAAGAATGA
AGCTGCTGGGTTGGAGGTGTCACGGCTCTGCTAGTGCTTGCTCTCCTCTTGGTGTG
CAGCTGGTCTGGATTGCTATGTCAAAGGTATGTGAAGGCCTCCCTTACAAACAAGAAT
CAGCAGAAGGAAATGATCGAAACCAAAAGTAGTAAAGGAGGAGAAGGCAATGATAGCAACCTAA
TGAGGAATCAAAGAAAATGATAAAACCCAGAAGAGTCCAAGAGTCCAAGCAAACCTACCGTGC
GATGCCTGGAAGCTGAAGTTTAGATGAGACAGAAATGAGGAGACACACTGAGGCTGGTTCTT
CATGCTCCTTACCCCTGCCCTGGGAAATCAAAGGGCAAAGAACCAAAGAACAGAATGCA
CCCTGGTTCTTAACGGATCAGCTCAGGACTGCCATTGGACTATGGAGTGCACCAAAG
GCCCTCTCCTTATTGTAACCTGTCTGGATCCATTCTCTACCTCCAAAGCTCCACGGCCT
TTCTAGCCTGGCTATGCTTAATAATATCCCACGGAGAAAGGAGTTGCAAAGTGCAAGGAC
CTAAAACATCTCATCAGTATCCAGTGGAAAAAGGCCTCTGGCTGTGAGGCTAGGTGGTTG
AAAGCCAAGGAGTCAGTGAGACCAAGGCTTCTACTGATTCCGCAGCTCAGACCCCTTCTCA
GCTCTGAAAGAGAACACGTATCCCACCTGACATGCTCTCTGAGCCCGTAAGAGCAAAGAAT
GGCAGAAAAGTTAGCCCTGAAAGCCATGGAGATTCTCATAACTTGAGACACTAATCTCTGTAAA
GCTAAAATAAGAAATAGAACAAAGGCTGAGGATACGACAGTACACTGTCAGCAGGGACTGTAAAC
ACAGACAGGGTCAAAGTGTCTCTGAACACATTGAGTTGAATCACTGTTAGAACACACACA
CTTACTTTCTGGTCTCTACCACTGCTGATATTTCTCTAGGAAATATACTTTACAAGTAACA
AAAATAAAACTCTTATAAATTCTATTGAGTTACAGAAATGATTACTAAGGAAGATT
ACTCAGTAATTGTTAAAAAGTAATAAAATTCAACAAACATTGCTGAATAGCTACTATATGTC
AAGTGCTGTGCAAGGTATTACACTCTGTAATTGAATATTATTCCTAAAAATTGCACATAGTAG
AACGCTATCTGGGAAGCTATTGTTCTAGTTGATATTCTAGCTTATCTACTTCCAAACTAAT
TTTATTGCTGAGACTAATCTTATTCTAATATGGCAACCATTATAACCTTAATT
TATTATTAACATACCTAAGAAGTACATGTTACCTCTATATACCAAAGCACATTAAAAGTGC
ATTAACAAATGTACTAGCCCTCTTTCCAACAAGAAGGGACTGAGAGATGCAGAAATATT
TGTGACAAAAATTAAAGCATTTAGAAAACCTT

FIGURE 6

MARCFSLVLLTSIWTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACRLLG
LSLAGKDQVETALKASFETCSYGWVGDFVVISRISPNSPKCGKNGVGVLIWKVPVSQFAAYCYN
SSDTWTNSCIPEIITKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPAPTTPPAPASTSI PR
RKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTALLVLALLFFGAAAGLGFCYVK
RYVKAFFPTNKNQQKEMIETKVVKEEKANDSNPNEESKKTDKNPEESKSPSKTTVRCLEAEV

Signal sequence:

amino acids 1-16

Transmembrane domain:

amino acids 235-254

N-glycosylation site.

amino acids 53-57, 130-134, 289-293

Casein kinase II phosphorylation site.

amino acids 145-149, 214-218

Tyrosine kinase phosphorylation site.

amino acids 79-88

N-myristoylation site.

amino acids 23-29, 65-71, 234-240, 235-239, 249-255, 253-259

FIGURE 7

CGCCCGCCTCCGCACCCGGCCCGCCACCGCGCCCTCCGCATCTGCACCCGAGCCCGC
GGCCTCCGGGGAGCGAGCAGATCCAGTCGGCCCGCAGCGCAACTCGGTCCAGTCGGGGCG
CGGCTCGGGCGCAGAGCGGAGATGCAGCGGCTGGGCCACCTGCTGTGCCTGCTGGCG
CGGCGGCCCCACGGCCCCCGCGCCCTCGACGGCACCTCGGTCCAGTCAGCAGCCGGCCCG
GCTCTCAGCTACCCGCAGGAGGCCACCTCAATGAGATGTTCCCGAGGTTGAGGAACGTGAT
GGAGGACACGCAGCACAAATTGCGCAGCGCGGTGGAAGAGATGGAGGCAGAAGAAGCTGCTGCTA
AAGCATCATCAGAAGTGAACCTGGCAAACCTACCTCCCAGCTATCACAATGAGACCAACACAGAC
ACGAAGGTTGAAATAATACCATCCATGTGCACCGAGAAATTACAAGATAACCAACAACCAGAC
TGGACAAATGGTCTTTCAGAGACAGTTACATCTGTGGGAGACGAAGAAGGCAGAAGGAGCC
ACGAGTGCATCATCGACGAGGACTGTGGGCCAGCATGTACTGCCAGTTGCCAGCTCCAGTAC
ACCTGCCAGCCATGCCGGGCCAGAGGATGCTCTGCACCCGGACAGTGAGTGTGTGGAGACCA
GCTGTGTCTGGGTCACTGCACCAAAATGCCACCAAGGGGAGCAATGGGACCATCTGTGACA
ACCAGAGGGACTGCCAGCCGGGCTGTGCTGTGCCCTCCAGAGAGGGCCTGCTGTTCCCTGTG
ACACCCCTGCCGTGGAGGGCAGCTTGCCATGACCCGCCAGCGGCTCTGGACCTCATCAC
CTGGGAGCTAGGCCATGGAGCCTGGACCGATGCCCTGTGCCAGTGGCCTCCTGTGCCAGC
CCCACAGCCACAGCCTGGTGTATGTGCAAGCCGACCTCGTGGGAGCCGTGACCAAGATGGG
GAGATCCTGCTGCCAGAGAGGTCCCCGATGAGTATGAAGTTGGCAGCTCATGGAGGAGGTGCG
CCAGGAGCTGGAGGACCTGGAGAGGAGCTGACTGAAGAGATGGCGTGGGGAGCCTGCGGCTG
CCGCCGCTGCACTGCTGGGAGGGAAAGAGATTAGATCTGGACCAGGCTGTGGTAGATGTGCAA
TAGAAATAGCTAATTATTCCCCAGGTGTGCTTAGGCGTGGCTGACCAGGCTTCCCTA
CATCTTCTCCAGTAAGTTCCCCTCTGGCTTGACAGCATGAGGTGTGCAATTGTTCAGCT
CCCCCAGGCTGTTCTCCAGGCTTCACAGTCTGGTGTGCTGGAGAGTCAGGCAGGGTAAACTGCA
GGAGCAGTTGCCACCCCTGTCCAGATTGGCTGCTTGCCCTACCAAGTGGCAGACAGCCG
TTTGTCTACATGGCTTGATAATTGTTGAGGGAGGAGATGAAACAATGTGGAGTCTCCCT
TGATTGGTTGGGAAATGTGGAGAAGAGTGCCTGCTTGCAAAACATCAACCTGGCAAAATG
CAACAAATGAATTTCACGCAGTTCTTCCATGGCATAGGTAAGCTGTGCTTCAGCTGTTGC
AGATGAAATGTTCTGTTCACCTGCATTACATGTGTTATTCACTCCAGCAGTGTGCTCAGCTCC
TACCTCTGTGCCAGGGCAGCATTTCATATCCAAGATCAATTCCCTCTCAGCACAGCCTGGG
AGGGGTATTGTTCTCGTCCATCAGGATCTCAGAGGCTCAGAGACTGCAAGCTGCTTGC
CAAGTCACACAGCTAGTGAAGACCAGAGCAGTTCATCTGGTTGTGACTCTAAGCTCAGTGC
CTCCACTACCCACACCAGCCTGGTGCCACAAAGTGTCTCCAAAAGGAAGGAGATGGGAT
TTTCTTGAGGCATGCACATCTGAAATTAGGTCAAACATTCTCACATCCCTCTAAAGTAAA
CTACTGTTAGGAACAGCAGTGTCTCACAGTGTGGGAGCCGTCTTAATGAAGACAATGAT
ATTGACACTGTCCCTTTGGCAGTTGCATTAGTAACTTGAAAGGTATATGACTGAGCGTAGCA
TACAGGTTAACCTGCAGAAACAGTACTTAGGTAATTGTAGGGCGAGGATTATAATGAAATTG
AAAATCACTAGCAGCAACTGAAGACAATTATCAACCACGTGGAGAAAATCAAACCGAGCAGGGC
TGTGTGAAACATGGTTGTAATATGCGACTGCGAACACTGAACCTACGCCACTCCACAAATGATG
TTTCAGGTGTATGGACTGTTGCCACCATGTATTCACTCAGAGTTCTTAAAGTTAAAGTTGCA
CATGATTGTATAAGCATGCTTCTTGAGTTAAATTATGTATAAACATAAGTTGCATTAGAA
ATCAAGCATAAAATCACTCAACTGCAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 8

MQRLGATLLCLLLAAAVPTAPAPAPTATSAPVKPGPALSYPQEEATLNEMFREVEELMEDTQHKL
RSAVEEMEAEEAAAKASSEVNLNLPPSYHNETNTDTKGNNTIHVHREIHKITNNQTGQMVFSE
TVITSGDEEGRSHECIIDEDCGPSMYCQFASFQYTCQPCRGQRLCTRSECCGDQLCVWGHCT
TKMATRGSNGTICDNQRDCQPGGLCAFQRGLFPVCTPLPVEGELCHDPASRLLDLITWELEPDG
ALDRCPCASGLLCQPHSHSLVYVCKPTFVGSRDQDGEILLPREVPDEYEVGSMEEVRQELEDLE
RSLTEEMALGEPAAAAALLGGEI

Signal sequence:

amino acids 1-19

N-glycosylation site.

amino acids 96-100, 106-110, 121-125, 204-208

Casein kinase II phosphorylation site.

amino acids 46-50, 67-71, 98-102, 135-139, 206-210, 312-316,
327-331

N-myristoylation site.

amino acids 202-208, 217-223

Amidation site.

amino acids 140-144

FIGURE 9

CGGACGCGTGGCGGGACGCGTGGGGCTGTGAGAAAGTCCAATAATAACATCATGCAACCCCAC
GCCACCTTGTGAACTCCTCGTGCCAGGGCTGATGTGCGTCTCCAGGGCTACTCATCAAAG
GCCTAATCCAACGTTCTGTCTCAATCTGCAAATCTATGGGGCCTGGGGCTTCTGGACCCCTT
AACTGGGTACTGGCCCTGGCCAATGCGTCCTCGCTGGAGCCTTGCCTCCTACTGGCCTT
CCACAAGCCCCAGGACATCCCTACCTTCCCTTAATCTGCCTCATCCGACACTCCGTTACC
ACACTGGGTATTGGCATTGGAGCCCTCATCCTGACCCCTGTGCAGATAGCCGGTCATCTG
GAGTATATTGACCACAAGCTCAGAGGAGTGCAGAACCTGTAGCCGCTGCATCATGTGCTGTT
CAAGTGCCTGCTGGTCTGGAAAAATTATCAAGTTCTAAACCGCAATGCATACATCATGA
TCGCCATCTACGGGAAGAATTCTGTCTCAGCCAAAATGCGTCATGCTACTCATGCGAAC
ATTGTCAGGGTGGTCGTCTGGACAAAGTCACAGACCTGCTGTTCTGGAAAGCTGCTGGT
GGTCGGAGGCGTGGGGCTGTCTTCTTCTCCGGTCGCATCCGGGGCTGGTAAAG
ACTTTAAGAGCCCCCACCTCAACTATTACTGGCTGCCATCATGACCTCCATCCTGGGGCCTAT
GTCATGCCAGCGCTTCTCAGCTTCCGGCATGTGTGGACACGCTCTCCTGCTTCT
GGAAGACCTGGAGCGAACACGGCTCCCTGGACCGCCCTACTACATGTCCAAGAGCCTCTAA
AGATTCTGGCAAGAAGAACGAGGCGCCCCGGACAACAAGAAGAGGAAGAAGTGACAGCTCCGG
CCCTGATCCAGGACTGCACCCACCCACCGTCCAGCCATCCAACCTCACCTGCCTTACAGGT
CTCCATTTGTGGTAAAAAAGGTTTAGGCCAGGCGCCGTGGCTCACGCCTGTAATCCAACACT
TTGAGAGGCTGAGGCGGGCGGATCACCTGAGTCAGGAGTCAGGACCGCCTGCCAACATGGT
AAACCTCCGTCTATTAAAAATAACAAAAATTAGCCAGAGTGGTGGCATGCACCTGTCATCCCA
GCTACTCGGAGGCTGAGGAGAATCGCTGAACCCGGAGGCAGAGGTTGCAGTGAGCCGA
GATCGGCCACTGCACTCCAACCTGGTGACAGACTCTGTCTCCAAAACAAACAAACAAA
AAGATTATTAAAGATATTGTTAACTC

FIGURE 10

RTRGRTRGGCEKVPINTSCNPTAHVNSSCPGLMCVFQGYSSKGLIQRSVFNLQIYGVLGLFWTL
NWVLALGQCVLAGAFASFYWAFHKPQDIPTFPLISAFIRTLRYHTGSLAFGALILTLVQIARVIL
EYIDHKLRGVQNPVARCIMCCFKCCLWCLEKFIKFLNRNAYIMIAIYGKNFCVSAKNAFMILLMRN
IVRVVVLDKVTDLLFFGKLLVVGVGVLSSFFSGRIPGLGKDFKSPHLNYWLPIMTSILGAY
VIASGFFSVFGMCVDTLFLCFILEDLERNNGSLDRPYYMSKSLKILGKKNEAPPDNKKRKK

Important features:

Transmembrane domains:

amino acids 57-80 (type II), 110-126, 215-231, 254-274

N-glycosylation sites.

amino acids 16-20, 27-31, 289-293

Hypothetical YBR002c family proteins.

amino acids 276-288

Ammonium transporters proteins.

amino acids 204-231

N-myristoylation sites.

amino acids 60-66, 78-84

Amidation site.

amino acids 306-310

FIGURE 11

GCCCCGCGCCCGCGCCGGCGCCGGCGCCGGAGCCACCGCCATGGGGCCTGCCTGGGAGCCTGC
TCCCTGCTCAGCTGCGCGTCTGCCTCTGGCTCTGCCCTGCATCCTGTGCAGCTGCTGCCCGC
CAGCCGCAACTCCACCGTGAGCCGCTCATCTCACGTTCTCCTCTCCTGGGGTGTGGAGGGGCC
TCATTATGCTGAGCCGGCGTGGAGAGTCAGCTCTACAAGCTGCCCTGGGTGTGGAGGGGCC
GGGATCCCCACCGTCTGCAGGCCACATCGACTGTGGCTCCCTGCTGGTACCGCGCTGTCTACCG
CATGTGCTTCGCCACGGCGCCTCTTCTTACCCCTGCTCATGCTCTGCGTGTGAGCAGCA
GCCGGGACCCCCGGGCTGCCATCCAGAATGGGTTGGTCTTAAGTTCTGATCCTGGTGGCCTC
ACCGTGGGTGCCTTCTACATCCCTGACGGCTCCTCACCAACATCTGGTTCTACTCGCGTCGTGGG
CTCCTCCTCTCATCCTCATCCAGCTGGTGTGCTCATCGACTTGCCTCCTGGAAACCAGCGGT
GGCTGGGCAAGGCCGAGGAGTGCATTCCGTGCCTGGTACGCAGGCCTCTTCTTCACTCTCCTC
TTCTACTTGCTGTCGATCGCGGCCGTGGCGTGTGATGTTCATGTACTACACTGAGCCCAGGGCTGCCA
CGAGGGCAAGGTCTTCATCAGCCTAACCTCACCTCTGTGTCTGCGTGTCCATCGCTGCTGTCTGC
CCAAGGTCCAGGACGCCAACCTGGGTCTGCTGCAGGCCTCGGTATCACCTTACACCATG
TTTGTACCTGGTCAGCCCTATCCAGTATCCCTGAACAGAAATGCAACCCCCATTGCCAACCCAGCT
GGCAACGAGACAGTTGTGGCAGGCCCGAGGGCTATGAGACCCAGTGGTGGATGCCCGAGCATTG
TGGGCCTCATCATCTCCTCTGTGCACCCCTTCACTCAGTCTGCGCTCCTCAGACCACCGCAGGTG
AACAGCCTGATGCAGACCGAGGAGTGCCACCTATGCTAGACGCCACACAGCAGCAGCAGCAGGT
GGCAGCCTGTGAGGCCGGGCTTGACAACGAGCAGGACGGCGTACCTACAGCTACTCCTTCTCC
ACTTCTGCCTGGTGTGGCTCACTGCACGTATGATGACGCTACCAACTGGTACAAGCCCAGTGG
ACCCGGAAGATGATCAGCACGTGGACCGCCGTGTGGTGAAGATCTGTGCCAGCTGGCAGGGCTGCT
CCTCTACCTGTGGACCCCTGGTAGCCCCACTCCTCTGCGAACCGCAGTCAGCTGGTCAAGCCCAGTGG
CAGCCTGCCATCTGGTGCCTCTGCCACCTGGTGCCTCTGGTACAGCCAACTGCCCTGAGCCGGC
CCCCACACCAATCAGCCAGGCTGAGCCCCACCCCTGCCAGCTCCAGGACCTGCCCTGAGCCGGC
CTTCTAGTCGTAGTGCCTCAGGGTCCGAGGAGCATCAGGCTCTGCGAGAGCCCCATCCCCCGCAC
ACCCACACGGTGGAGCTGCCTCTCCTCCCTCCCTGTTGCCATACTCAGCATCTGGATGAA
AGGGCTCCCTGTCCCTCAGGCTCCACGGGAGCGGGCTGCTGGAGAGAGCGGGAACTCCCACCAAG
TGGGCATCCGGCACTGAAGCCCTGGTGTGCTGGTACGTCCCCCAGGGACCCCTGCCCTTCCTG
GACTCGTGCCTTACTGAGTCTAAGACTTTCTAATAAACAGCCAGTGCAGTGTAAAAAAA

FIGURE 12

MGACLGACSLSCASCLCGSAPCILCSCCPASRNSTVSRLIFTFFLFLGVLVSIIMLSPGVESQL
YKLPWVCEEGAGIPTVLQGHIDCGSLLGYRAVYRMCFATAAFFFTLLMLCVSSSRDPRAAIQ
NGFWFFKFLILVGLTVGAFYIPDGSFTNIWFYFGVVGSLFILQLVLLIDFAHWNQRWLGKAE
ECDSRAWYAGLFFFLLFYLLSIAAVALMFYYTEPSGCHEGVFISLNLTFCVCVSIAAVLPKV
QDAQPNSGLLQASVITLYTMFVTSALSSIPEQKCNPHLPTQLGNETVVAGPEGYETQWWDAPSI
VGLIIFLLCTLFISLRSSDHRQVNSLMQTEECPPMLDATQQQQQQVAACEGRAFDNEQDGVTYSY
SFFHFCLVLASLHVMMTLTNWYKPGTRKMISTWTAVWVKICASWAGLLLWTLVAPLLLNRND
FS

Signal sequence:

amino acids 1-20

Transmembrane domains:

amino acids 40-58, 101-116, 134-150, 162-178, 206-223, 240-257,
272-283, 324-340, 391-406, 428-444

FIGURE 13

CGGGCCAGCCTGGGGCGGCCAGGAACCACCGTTAAGGTGTCTCTTTAGGGATGGTGA
GGTTGGAAAAAGACTCCTGTAACCCTCCTCCAGGATGAACCACCTGCCAGAAGACATGGAGAACG
CTCTCACCGGGAGCCAGAGCTCCCATGCTCTGCGCAATATCCATTCCATCAACCCCACACAA
CTCATGCCAGGATTGAGTCCTATGAAGGAAGGGAAAAGAAAGGCATATCTGATGTCAGGAGGAC
TTTCTGTTGTTGTCACCTTGACCTTATTGTAACATTACTGTGGATAATAGAGTTAAATG
TGAATGGAGGCATTGAGAACACATTAGAGAAGGAGGTGATGCAGTATGACTACTATTCTCATAT
TTTGATATATTCTTCTGGCAGTTTCGATTAAAGTGTAAATACCTGCATATGCTGTGCGAG
ACTGCGCCATTGGTGGCAATAGCGTTGACAACGGCAGTGACCAGTGCCTTTACTAGCAAAAG
TGATCCTTCGAAGCTTCTCAAGGGCTTGGCTATGTGCTGCCATCATTCAATTCA
CTTGCCTGGATTGAGACGTGGTCCTGGATTCAAAGTGTACCTCAAGAACAGAAGAAGAAAA
CAGACTCCTGATAGTCAGGATGCTCAGAGAGGGCAGCACTTACCTGGTGTCTGATG
GTCAGTTTATTCCCTCCTGAATCCGAAGCAGGATCTGAAGAACGCTGAAGAAAAACAGGACAGT
GAGAAACCACTTAGAAACTATGAGTACTACTTTGTTAAATGTGAAAAACCTCACAGAAAGTC
ATCGAGGCACAAAGAGGCAGGCAGTGAGTCTCCGTGACAGTAAAGTTGAAATGGTACGTC
CACTGCTGGCTTATTGAACAGCTAATAAAGATTATTGTAATACCTCACAAACGTTGTAC
CATATCCATGCACATTAGTCAGGCTGCTGGCTGGTAAGGTAATGTCATGATTCAATTCTCT
TCAGTGAGACTGAGCCTGATGTGTTAACAAATAGGTGAAGAAAGTCTGTGCTGTATTCTTAATC
AAAAGACTTAATATTGAAGTAACACTTTAGTAAGCAAGATACTTTATTCAATTCA
AGAATGGAATTTTTGTTCATGTCTCAGATTATTTGTTCTTTAACACTCACATT
TCCCTGTTTTAACTCATGCACATGTGCTTTGTCAGTTAAAAAGTGTAAATAAAATCTG
ACATGTCAATGTGGCTAGTTTATTCTGTTGCATTATGTGTATGGCCTGAAGTGTGGA
CTTGCAAAAGGGAAAGAAAGGAATTGCGAATACATGTAAGGACATTCAGACATTGTATTATT
TTATCATGAAATCATGTTCTGATTGTTCTGAAATGTTCTAAATACTCTTATTGAAATGC
ACAAAATGACTTAAACCATTCAATCATGTTCCCTTGCAGCCAATTCAATTAAAATGAA
CTAAATTAAAAA

FIGURE 14

MNHLPEDMENALTGSQSSHASLRNIHSINPTQLMARIESTYEGREKKGISDVRRTFCLFVTFDLLF
VTLLWIIELNVNGGIENTLEKEVMQYDYYSSYFDIFLLAVFRFKVLILAYAVCRLRHWWAIALTT
AVTSAFLLAKVILSKLFSQGAFGYVLPIISFILAWIETWFLDFKVLQPQEAEENRLLIVQDASER
AALIPGGLSDGQFYSPPESEAGSEEAEEKQDSEKPLLEL

Important features of the protein:

Signal peptide:

amino acids 1-20

Transmembrane domains:

amino acids 54-72, 100-118, 130-144, 146-166

N-myristoylation sites.

amino acids 14-20, 78-84, 79-85, 202-208, 217-223

FIGURE 15

ACTCGAACGCAGTTGCTTCGGGACCCAGGACCCCTCGGGCCGCCAGGAAAGACTGAGG
CCGGGGCCTGCCCGCCGGCTCCCTGCGCCGCCGCTCCCGGGACAGAAGATGTGCTCCAG
GGTCCCTCTGCTGCTGCCGCTGCTCTGCTACTGGCCCTGGGGCCTGGGTGCAAGGGCTGCCAT
CCGGCTGCCAGTGCAGCCACAGACAGTCTTCTGCACTGCCGCCAGGGACACGGTGCC
CGAGACGTGCCACCCGACACGGTGGGGCTGTACGTCTTGAGAACGGCATACCATGCTGACGC
AGGCAGCTTGCCGGCCTGCCGGCCTGCAAGCTCTGGACCTGTACAGAACCCAGATGCCAGCC
TGCCCAGCGGGTCTTCCAGCCACTGCCAACCTCAGCAACCTGGACCTGACGCCAACAGGCTG
CATGAAATACCAATGAGACCTTCCGTGGCCTGCGGCCCTGAGCGCCTTACCTGGCAAGAA
CCGCATCCGCCACATCCAGCCTGGTGCCTCGACACGCTCGACCCTCCTGGAGCTCAAGCTGC
AGGACAACGAGCTGCCGGCACTGCCCGCTGCCCTGCCGCCCTGCTGCTGCTGGACCTCAGC
CACACAGCCTCTGCCCTGGAGCCGGCATCTGGACACTGCCAACGTGGAGGCCCTGCC
GGCTGGTCTGGGGCTGCAGCAGCTGGACGAGGGCTTCAAGCCGCTTGCACGCCAACCTCCACGACC
TGGATGTGTCGACAACCAGCTGGAGCGAGTGCCACCTGTGATCCGAGGCCCTCGGGGCTGACG
CGCCTGCGGCTGCCGGCAACACCCGATTGCCAGCTGCCGCCAGGACCTGCCGGCCTGGC
TGCCTGCAAGGAGCTGGATGTGAGAACCTAACGCTGCAGGCCCTGCCTGGGACACTGCC
TCTTCCCCCGCTGCCGTGCTGGCAGCTGCCCAACCCCTCAACTGCGTGTGCCCTGAGC
TGGTTGGCCCTGGGTGCGCGAGAGCCACGTACACTGCCAGCCCTGAGGAGACGCCGTGCCA
CTTCCGCCCAAGAACGCTGCCGGCTGCTGGAGCTGACTACGCCACTTGGCTGCCAG
CCACCACCACACAGCCACAGTGCCACACGCCAGGGCCGTGGTGCGGGAGGCCACAGCCTGTCT
TCTAGCTGGCTCTACCTGGCTAGCCCCACAGCGCCACTGAGGCCAGGCCACCTGCC
CACTGCCCAACCGACTGTAGGGCCTGCCCCAGGCCAGGACTGCCACCGTCCACCTGCC
ATGGGGCACATGCCACCTGGGACACGGCACCACTGGCGTGTGCCCCGAAGGTTACG
GGCCTGTACTGTGAGAGCCAGATGGGCAGGGACACGCCAGCCCTACACCAGTCACGCCAG
GCCACCACGGTCCCTGACCCCTGGCATCGAGCCGGTGAGGCCACCTCCCTGCGGTGGGCTGC
AGCGCTACCTCCAGGGAGCTCCGTGCAGCTCAGGAGCCCTCGTCACTATGCCAACCTATCG
GGCCCTGATAAGCGGTGGTGACGCTGCCACTGCCCTGCCCTCGCTGAGTACACGGTACCCA
GCTGCCGCCAACGCCACTTACTCCGTGTGTCATGCCCTGGGCCGGGCGGGTGCCTGGGAGG
GCGAGGAGGCCTGCCGGAGGCCATACACCCAGCCGTCCACTCCAACCACGCCAGTCACC
CAGGCCCGCAGGGCAACCTGCCGTCCATTGCCGCCCTGCCCGCGGTGCTCTGCC
GCTGGCTGCCGTGGGGCAGCTACTGTGTCGGCGGGGGCGGCATGGCAGCAGGGCTCAGG
ACAAAGGGCAGGTGGGGCAGGGCTGGGCCCTGGAACACTGGAGGGAGTGAAGGTCCC
CCAGGCCCGAAGGCACAGAGGGCGGTGGAGAGGCCCTGCCAGCGGTCTGAGTGTGAGGTGCC
ACTCATGGGCTTCCAGGGCTGGCTCCAGTCACCCCTCCAGCAAAAGCCCTACATTAAGCCA
GAGAGAGACAGGGCAGCTGGGCCGGCTCTCAGGCCAGTGAAGATGCCAGGCCCTCC
ACACCACGTAAGTTCTCAGTCCAAACCTGGGATGTGTCAGACAGGGCTGTGACCAAGCT
GGGCCCTGTTCCCTCTGGACCTCGGTCTCCTCATCTGTGAGATGCTGTGGGCCAGCTGACGCC
CTAACGTCCCCAGAACCGAGTGCCTATGAGGACAGTGTCCGCCCTGCCCTCGAACGTGCAGTC
CCTGGCACGGCGGGCCCTGCCATGTGTTGTAACGCATGCCCTGGGTCTGCTGGCTCTCC
TCCAGGCAGGACCTGGGGCCAGTGAAGGAAGCTCCGGAAAGAGCAGAGGGAGAGCGGGTAGGC
GGCTGTGTGACTCTAGTCTTGGCCCCAGGAAGGAAGGAACAAAGAAACTGGAAAGGAAGATGC
TTTAGGAACATGTTGCTTTAAATATATATTTATAAGAGATCCTTCCCATTATTCT
GGGAAGATGTTTCAAACTCAGAGACAAGGACTTGGTTTGTAAAGACAAACGATGATATGAA
GCCCTTTGTAAGAAAAATAAAAGATGAAGTGTGAAA

FIGURE 16

MCSRVPLLLPLLLLALGPGVQGCPSCQCSQPQTVFCTARQGTTVPRDVPPDTVGLYVFENGIT
MLDAGSFAGLPGLQLLDLSQNQIASLPSGVFQPLANLSNLDLTANRLHEITNETFRGLRRLERLY
LGKNRIRHIQPGAFDTLDRLLELKLQDNELRALPPLRLPRLLLLDLHSNSLLALEPGILDTANVE
ALRLAGLGLQQLDEGLFSRLRNLHLDVSDNQLERVPPVIRGLRGLTRLRLAGNTRIAQLRPEDL
AGLAALQELDVSNLSLQALPGDLSGLFPRLRLLAAARNPFNCVCPLSWFGPWVRESHVTLASPEE
TRCHFPPKNAGRLLLELDYADFGCPATTTATVPTTRPVVREPTALSSSLAPTWLSPTAPATEAP
SPPSTAPPTVGPVPQPQDCPPSTCLNGGTCHLGTRHHLACLCPEGFTGLYCESQMGQGTRPSPTP
VTPRPPRSLTLGIEPVSPSLRVGLQRYLQGSSVQLRSRLTYRNLSGPDKRLVTLRLPASLAEY
TVTQLRPNATYSVCVMPILGPGRVPEGEAACGEAHTPPAVHSNHAPVTQAREGNLPLLIAPALAAV
LLAALAAVGAAYCVRRGRAMAAAQDKGQVGPAGPLELEGVKVPLEPGPKATEGGEALPSGSE
CEVPLIMGFPGPGLQSPHLAKPYI

Important features:

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 579-599

EGF-like domain cysteine pattern signature.

amino acids 430-442

Leucine zipper pattern.

amino acids 197-219, 269-291

N-glycosylation sites.

amino acids 101-105, 117-121, 273-277, 500-504, 528-532

Tyrosine kinase phosphorylation sites.

amino acids 124-131, 337-345

N-myristoylation sites.

amino acids 23-29, 27-33, 70-76, 142-148, 187-193, 348-354,
594-600, 640-646

FIGURE 17

GCAGCGGCGAGGCAGGCGGTGGCTGAGTCGTGGCAGAGCGAAGCGACAGCTCATGCG
GGTCCGGATAGGGCTGACGCTGCTGCTGTGCGGTGCTGAGCTGGCCTCGCGTCCTCGG
ATGAAGAAGGCAGCCAGGATGAATCCTAGATTCCAAGACTACTTGACATCAGATGAGTCAGTA
AAGGACCATACTACTGCAGGCAGAGTAGTTGCTGGTCAAATATTCTTGATTCAAAGAATCTGA
ATTAGAATCCTCTATTCAAGAAGAGGAAGACAGCCTCAAGAGCCAAGAGGGGAAAGTGTACAG
AAGATATCAGCTTCTAGAGTCTCAAATCCAGAAAACAAGGACTATGAAGAGCCAAGAAAGTA
CGGAAACCAGCTTGACCGCCATTGAAGGCACAGCACATGGGAGCCCTGCCACTCCCTTTCT
TTTCCTAGATAAGGAGTATGATGAATGTACATCAGATGGGAGGGAAAGATGGCAGACTGTGGTGTG
CTACAACCTATGACTACAAAGCAGATGAAAAGTGGGCTTTGTGAAACTGAAGAAGAGGCTGCT
AAGAGACGGCAGATGCAGGAAGCAGAAATGATGTATCAAACGGAAATGAAATCCTTAATGGAAG
CAATAAGAAAAGCCAAAAAAGAGAAGACATATCGGTATCTCCAAAGGCAGCAAGCATGAACCATA
CCAAAGCCCTGGAGAGAGTGTATGCTCTTTATTGGTGAATTACTGCCACAGAATATCCAG
GCAGCGAGAGAGATGTTGAGAAGCTGACTGAGGAAGGCTCTCCAAGGGACAGACTGCTTTGG
CTTCTGTATGCCTCTGGACTTGGTGTAAATTCAAGTCAGGCAAAGGCTTGTATATTACAT
TTGGAGCTCTGGGGCAATCTAATAGCCCACATGGTTGGTAAGTAGACTTAGTGGAAAGGCT
AATAATATTAACATCAGAAGAATTGGTTATAGCGGCCACAACCTTTCAGCTTCATGATC
CAGATTGCTTGTATTAAGACCAAATTCAGTTGAACCTCCTCAAATTCTGTAAATGGATAT
AACACATGGAATCTACATGAAATGAAAGTTGGTGGAGTCCACAATTTCCTTAAATGATTAG
TTTGGCTGATTGCCCTAAAAGAGAGATCTGATAAATGGCTTTAAATTTCTCTGAGTTG
GAATTGTCAGAATCATTTCATGATTACATTGATTATCATAATTAAAAATTTCCTTAAATGGATAT
AAATTGTAATGGTGGCTATAGAAAACAACATGAAATATTACAATATTGCAACAATGC
CCTAAGAATTGTTAAAATTGAGTTATTGTCAGAATGACTCCAGAGAGCTACTTCTG
TTTTTACTTTCATGATTGGCTGTCTCCCTTTATTCTGGTCAATTGCTACTTTCTTGTAAATTGG
GCCTGCTCCAGTAGTCTCATTCCCTTTGCTAATTGTTACTTTCTTGTAAATTGG
AAGATTAACTCATTTAATAAAATTATGCTAAGATTAAGATTAAGATTAAGATTAAGATTAAG
AAA

FIGURE 18

MRVRIGLTLLCAVLLSLASASSDEEGSQDES LDSKTTLSDESVKDHTTAGRVVAGQIFLDSEESEL
ESSIQEEEDSLKSQEGERVTEDISFLESPNPNKDYEEPKKVRKPALTAIEGTAHGEPCFPFLFLDK
EYDECTS DGDREDGRLWCATTYDYKADEKWGFCETEEAAKRRQMQEAMMYQTGMKILNGSNKKSQKR
EAYRYLQKAASMNHTKALERVSYALLFGDYLQPQNIQAAREMFEKLTEEGSPKGQTALGFLYASGLGVN
SSQAKALVYYTFGALGGNLIAHMVLVSRL

Important features:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 195-199, 217-221, 272-276

Tyrosine kinase phosphorylation site.

amino acids 220-228

N-myristoylation sites.

amino acids 120-126, 253-259, 268-274, 270-274, 285-291, 289-295

Glycosaminoglycan attachment site.

amino acids 267-271

Microbodies C-terminal targeting signal.

amino acids 299-303

Type II fibronectin collagen-binding domain protein.

amino acids 127-169

Fructose-bisphosphate aldolase class-II protein.

amino acids 101-119

FIGURE 19

AATTCAGATTAAAGCCCATTCTGCAGTGGAAATTCACTGAACACTAGCAAGAGGACACCATCTTCTT
GTATTATAACAAGAAAGGAGTGTACCTATCACACACAGGGGGAAAAATGCTTTGGGTGCTAGG
CCTCCTAATCCTCTGTGGTTCTGTGGACTCGTAAAGGAAACTAAAGATTGAAGACATCACTG
ATAAGTACATTTTATCACTGGATGTGACTCGGCTTGAAACTGGCAGCCAGAACTTTGAT
AAAAAGGGATTCATGTAATCGCTGCCTGTACTGAATCAGGATCAACAGCTTAAAGGAGA
AACCTCAGAGAGACTTCGTACTGTGCTTCTGGATGTGACCGACCCAGAGAATGTCAAGAGGACTG
CCCAGTGGGTGAAGAACCAAGTTGGGGAGAAAGGTCTCTGGGTCTGATCAATAATGCTGGTGT
CCCAGCGTGCTGGCTCCACTGACTGGCTGACACTAGAGGACTACAGAGAACCTATTGAAGTGA
CCTGTTGGACTCATCAGTGTGACACTAAATATGCTTCCTTGGTCAAGAAAGCTCAAGGGAGAG
TTATTAATGTCAGTGGAGGTGCCTGCAATCGTGGAGGGCTATACTCCATCCAAA
TATGCAGTGGAAAGGTTCAATGACAGCTTAAGACGGGACATGAAAGCTTTGGTGTGCACGTCTC
ATGCATTGAACCAGGATTGTTCAAAACAAACTGGCAGATCCAGTAAAGGTAATTGAAAAAAAC
TCGCCATTGGGAGCAGCTGTCCAGACATCAAACAACATATGGAGAAGGTTACATTGAAAAA
AGTCTAGACAAACTGAAAGGCAATAATCCTATGTGAACATGGACCTCTCCGGTAGAGTG
CATGGACCACGCTCTAACAGTCTTCCCTAACAGACTCATTATGCCGCTGGAAAGATGCCAAA
TTTCTGGATACCTCTGTCTCACATGCCAGCTTGCAAGACTTTATTGTTGAAACAGAAA
GCAGAGCTGGCTAATCCCAAGGCAGTGTGACTCAGCTAACCAACATGTCTCCAGGCTATGA
AATTGGCGATTCAAGAACACATCTCCTTTCAACCCATTCTTATCTGCTCCAACCTGGACT
CATTTAGATCGTGTATTGGATTGCAAAAGGGAGTCCACATCGCTGGTGTATCCCAGGGT
CCCTGCTCAAGTTCTTGAAAAGGAGGGCTGGAATGGTACATCACATAGGCAAGTCCTGCCCT
GTATTAGGCTTGCCTGCTGGTGTGATGTAAGGAAATTGAAAGACTTGCCATTCAAATGA
TCTTACCGTGGCCTGCCCATGCTTATGGTCCCCAGCATTACAGTAACGTGAATGTTAAGT
ATCATCTCTTATCTAAATATTAAAGATAAGTCAACCCAAAAAA
AAAAAA

FIGURE 20

MLFWVLGLLILCGFLWTRKGKLKIEDITDKYIFITGCDSGFGNLAARTFDKKGFHVIACLTESG
STALKAETSERLRTVLLDVTDPENVKRTAQWVKNQVGEKGLWGLINNAGVPGVLAPTDWLTLEDY
REPIEVNLFGGLISVTLNMLPLVKKAQGRVINVSSVGGRLAIVGGGYTPSKYAVEGFNDSLRRDMK
AFGVHVSCIEPGLFKTNLADPVKVIEKKLAIWEQLSPDIKQQYGEFYIEKSLDKLGNKSYVNMD
LSPVVECMDHALTSLFPKTHYAGKDAKIFWIPLSHMPAALQDFLLLKQKAE LANPKAV

Important features of the protein:

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 136-152

N-glycosylation sites.

amino acids 161-163, 187-190 and 253-256

Glycosaminoglycan attachment site.

amino acids 39-42

N-myristoylation sites.

amino acids 36-41, 42-47, 108-113, 166-171, 198-203 and 207-212

FIGURE 21

CTGAGGC GGCGGTAGCATGGAGGGGGAGAGTACGT CGGCGGTGCTCGGGCTTGCTCGGCG
CACTCGCTTCCAGCACCTAACACGGACTCGGACACGGAAAGGTTCTTCTGGGAAGTAAA
GGTGAAGCCAAGAACAGCATTACTGATTCCAAATGGATGATGTTGAAGTTGTTATAACATTGA
CATTCA GAAATATATTCCATGCTATCAGCTTTAGCTTTATAATTCTCAGGCGAAGTAAATG
AGCAAGCACTGAAGAAAATATTATCAAATGTCAAAAGAATGTGGTAGGTTGGTACAAATTCCGT
CGTCATTCA GATCAGATCATGACGTTAGAGAGAGGCTGCTCACAAAAC TTGCAGGAGCATT
TTCAAACCAAGACCTGTTCTGCTATTAAACACCAAGTATAAACAGAAAGCTGCTACTC
ATCGACTGGAACATTCCCTATATAAACCTCAAAAGGACTTTCACAGGGTACCTTAGTGGTT
GCCAATCTGGCATGCTGAACA ACTGGTTATAAAACTGTATCAGGTTCTGTATGCCACTGG
TTTAGCCGAGCAGTACAAACACACAGCTCTAAATTTGAAGAAGATGGATCCTAAAGGAGG
TACATAAGATAAAATGAAATGTATGCTTCA TTACAAGAGGAATTAAAGAGTATATGCAAAAAGTG
GAAGACAGTGAACAAGCAGTAGATAACTAGTAAAGGATGTAACAGATTAAAACGAGAAATTGA
GAAAAGGAGAGGAGCACAGATTCAAGGAGCAAGAGAGAAGAACATCCAAAAGACCCTCAGGAGA
ACATTTCTTCAGGCATTACGGACCTTTCCAATTCTGAATTCTCATTCA GTGTT
ATGCTTTAAAAAATAGACATGTTCTAAAGTAGCTGTA ACTACAACCACATCTCGATGTAGT
AGACAATCTGACCTTAATGGTAGAACACACTGACATTCTGAAGCTAGTCCAGCTAGTACACCAC
AAATCATTAAAGCATAAAGCCTAGACTTAGATGACAGATGGCAATTCAAGAGATCTGGTTGTTA
GATAACACAAGACAAACGATCTAAAGCAAATCTGGTAGTAGTAACCAAGATAAGCATCCAAAAT
GAGCAGCCCAGAAACAGATGAAGAAATTGAAAAGATGAAGGGTTGGTGAATATTCACGGTCTC
CTACATTTTGATCCCTTAACCTACAAGGAGATTTTATTGGCTGATGGTAAAGCCAAAC
ATTTCTATTGTTTACTATGTTGAGCTACTGCAGTAAGTCATTGTTACTATGTCACC
TGTTGCAGTAATACACAGATAACTCTTAGTGCATTACTCACAAAGTACTTTCAAACATCA
GATGCTTTATTCCAACCTTTTACCTTCACTAAGTTGAGGGAAAGGCTTACACAG
ACACATTCTTAGAATTGAAAAGTGAGACCAGGCACAGTGGCTCACACCTGTAATCCCAGCACT
TAGGGAAGACAAGTCAGGAGGATTGATTGAAGCTAGGAGTTAGAGACCAGCCTGGCAACGTATT
GAGACCAGTCTATTAAAAAATGAAAAGCAAGAATAGCCTATTTCAAAATATGGAAA
GAAATTATGAAAATTCTGAGTCATTAAATTCTCCTTAAGTGATACTTTAGAAGTA
CATTATGGCTAGAGTTGCCAGATAAAATGCTGGATATCATGCAATAAATTGCAAAACATCATCT
AAAATTAAAAAAAAAAAAAAA

FIGURE 22

MEGESTSAVLSGFVLGALAFQHLNTSDTEGFLGEVKGEAKNSITDSQMDDVEVVTIDIQKYI
PCYQLFSFYNSSGEVNEQALKKILSNVKKNVVGWYKFRRHSDQIMTFRERLLHKNLQEHFSNQDL
VFLLLTPSIITESCSTHRLEHSLYKPQKGLFHRVPLVVANLGMSEQLGYKTVSGSCMSTGFSRAV
QTHSSKFFED GSLKEVHKINEMYASLQEELKSICKVEDSEQAVDKLVKDVRNLKREIEKRRGA
QIQAAREKNIQKDPQENIFLCQALRTFFPNSEFLHSCVMSLKNRHSKSSCNHHLDVVDNLTL
MVEHTDIPEASPASTPQIIKHKALDLDDRWQFKRSRLLDTQDKRSKANTGSSNQDKASKMSSPET
DEEIEKMKGFGEYSRSPTF

Important features:

Signal peptide:

amino acids 1-19

N-glycosylation sites.

amino acids 75-79, 322-326

N-myristoylation site.

amino acids 184-154

Growth factor and cytokines receptors family.

amino acids 134-150

FIGURE 23

GGCACAGCCGCGCGCGGGAGGGCAGAGTCAGCCGAGCCGAGTCCAGCCGGACGAGCGGACCAAGCGCAGGGCAGCCCAA
GCAGCGCGCAGCGAACGCCGCCGCCACACCCCTCTGCGTCCCCGGCGCCTGCCACCCCTCCCTCCCTCCCC
GCGTCCCCGCCCTGCCGGCCAGTCAGCTTGCCGGGTCGCTGCCCGCGAAACCCGAGGTACCCAGCCGCCCT
GCTTCCCTGGGCCGCGCGCCCTCCACGCCCTCCTCTCCCTGGCCCGCGCCTGGCACCGGGGACCGTTGCCTGA
CGCGAGGCCAGCTACTTTGCCCGCGTCCCTGCCCTGCTGCCCTTCCACCAACTCCAACCTCCCTCTCCC
TCCAGCTCCACTCGCTAGTCCCAGCTCCGCCAGCCCTGCCGCTGCCGTAGCGCCGCTTCCCGTCCGGTCCCAA
GGTGGGAACCGCGTCCGCCCGCCGACCAATGGCACGGTTCGGCTGCCCGCCTCTGCACCCGGCAGTGCTC
AGCGCCGCGCTGGCTGCCGAGCTAAGTCGAAAGTTGCTCGGAAGTGCAGCTTTACGTGTCAAAGGCTTC
AACAAAGAACGATGCCCTCCACGAGATCAACGGTATCATTGAAGATCTGTCCCAGGGTCTACCTGCTGCTCT
CAAGAGATGGAGGAGAAGTACAGCCTGCAAAGTAAAGATGATTTCAAAGTGTGGCAGCGAACAGTGCAATCATTG
CAAGCTGTCTTGCTTCACTAACAGAAGTTGATGAATTCTCAAAGAAACTACTGAAAATGCAGAGAAATCCCTG
AATGATATGTTGTGAAGACATATGCCATTATACATGAAAATTCTGAGCTATTAAAGATCTTCGTAGAGTTG
AAACGTTACTACGTGGTGGAAATGTGAACCTGAAAGAAATGCTAAATGACTTCTGGCTGCCCTGGAGCGGATG
TTCCGCCCTGGTGAACCTCCAGTACCACTTACAGATGAGTATCTGGAATGTGTGAGCAAGTACGGAGCAGCTGAAG
CCCTCGGAGATGTCCCTCGCAAATTGAAGCTCCAGGTTACTCGTGTCTTGAGCAGCCGACTTCGCTCAAGGC
TTAGCGGTTGCGGGAGATGTCGTGAGCAAGGTCTCCGTGGTAAACCCACAGCCAGTGTACCCATGCCCTGTTGAAG
ATGATCTACTGCTCCACTGCCGGGTCTCGTACTGTGAAGCCATGTTACAACACTGCTCAAACATCATGAGAGGC
TGTTGGCCAACCAAGGGATCTGATTGAATGAAACAATTCTAGATGCTATGCTGATGGTGGCAGAGAGGCTA
GAGGGCCTTCAACATTGAATCGGTATGGATCCACGAGGTTACTGAGCAAGTGGACACCAGCTGGACGAATTCTCGT
AATAGTGTCAAGTGTCTCAGAAGGTTCCAGGGATGTGGACCCCCAAGCCCTCCAGCTGGACGAATTCTCGT
TCCATCTCTGAAAGTGCCTCAGTGCCTCAGACCACATCACCCGAGGAACGCCAACACAGCAGCTGGCACT
AGTTGGACCGACTGGTTACTGATGTCAGGAGAAACTGAAACAGGCAAGAAATTCTGGCCTCCCTCCGAGCAAC
GTTTGCACGATGAGAGGATGGCTGCAGGAAACGGCAATGAGGATGACTGTTGGAAATGGAAAGGCAAAAGCAGGTAC
CTGTTGCAGTGACAGGAAATGGATTAGCCAACCGGGCAACAACCCAGAGGTCCAGGTTGACACCAGCAAACCGAGAC
ATACTGATCCTCGTCAAATCATGGCTCTCGAGTGTGATGACCAGCAAGATGAAGAATGCATAACATGGAACGACGTG
GACTTCTTGATATCAGTGTGAAAGTAGTGGAGAAGGAAGTGGAGTGGCTGTGAGTATCAGCAGTGCCCTCAGAG
TTTGAATGCAATGCCACTGACCATGCTGGGAAGAGTGCCTGAAATGAGAAAGCCGACAGTGCTGGTGTCCGCTGGGCA
CAGGCCTACCTCCTCACTGCTTCTGCATCTGTTCTGGTTATGCAGAGAGAGTGGAGAATATTCTCAAACCTCTGAG
AAAAAGTGTCAAAAGTTAAAGGCACCAAGGTTACCTTACCATCCTAGTGACTTGCTTTCTCATTCAAGTTGGG
TGGACAACAATGTACAGTTTACTATGTGGCACTGGTTAAGAAGTGCTGACTTGTTCTCATTCAAGTTGGG
AGGAAAAGGGACTGTGCATTGAGTGGTCTGCTCCCCAAACCATGTTAACAGTGGCTAACAGTGAGGTACAGAA
CTATAGTTAGTGTGCATTGATGTTCTTACACTCTATTATGTTGATGTTTTCTCATTGTTGTGGGTT
TTTTTCCAACTGTGATCTGCCTGTTACAAGCAAACCAAGGGTCCCTTGGCACGTAACATGTACGTATT
TCTGAAATATTAAATAGCTGTACAGAAGCAGGTTTATTATCATGTTATCTTATTAAAAGAAAAAGCCAAAAGC

FIGURE 24

MARFGLPALLCTLAVLSAALLAAELKS KSCSEV RRLYVSKGF NKDAPLHEINGDHLKICPQGST
CCSQEMEEKYSLQSKDDFKSVVSEQCNHLQAVFASRYKKFDEFFKELLENAEKSLNDMFVKTYGH
LYMQNSELFKDLFVELKRYYVVG NVNLEEMLNDFWARLLERMFR LVNSQYHFTDEYLECVSKYTE
QLKPF GDVPRKLKLQVTRAFVAARTFAQGLAVAGDVSKVSVNPTAQCTHALLKMIYCSHC RGL
VTVKPCYN YCSNIMRGCLANQGDLD FEWNNFIDAMLMVAERLEGPFNIESVMDPIDVKISDAIMN
MQDNSVQVSQKV FQGCGPPKPLPAGRISRSI SEAFSARFRPHPEERPTTAAGTSLDRLVTDVK
EKLKQAKKFWSSLPSNCNDERMAAGNGNEDDCWNGKGKS RYLFAVTGNGLANQGNP EVQV DTS
KPDILILRQIMALRVMTSKMKNAYNGNDV DFFDISDESSGEGSGSGCEYQQCPSEFDYNATDHAG
KSANEKADSAGVRPGAQAYLLTVFCILFLVMQREWR

Important features:

Signal peptide:

amino acids 1-22

ATP/GTP-binding site motif A (P-loop).

amino acids 515-524

N-glycosylation site.

amino acids 514-518

Glycosaminoglycan attachment sites.

amino acids 494-498, 498-502

N-myristoylation sites.

amino acids 63-69, 224-230, 276-282, 438-444, 497-503, 531-537

Glypicans proteins.

amino acids 54-75, 105-157, 238-280, 309-346, 423-460, 468-506

FIGURE 25

CTCGCCCTCAAATGGGAACGCTGGCCTGGACTAAAGCATAGACCACCAGGCTGAGTATCCTGAC
CTGAGTCATCCCCAGGGATCAGGAGCCTCCAGCAGGGAACCTCCATTATATTCTTCAAGCAACT
TACAGCTGCACCGACAGTTGCGATGAAAGTTCTAATCTCTCCCTCCTGTTGCTGCCACTAA
TGCTGATGTCCATGGTCTCTAGCAGCCTGAATCCAGGGGTCGCCAGAGGCCACAGGGACCGAGGC
CAGGCTTCTAGGAGATGGCTCCAGGAAGGCGGCCAAGAATGTGAGTGCAAAGATTGGTTCTGAG
AGCCCCGAGAAGAAAATTCATGACAGTGTCTGGCTGCCAAAGAACAGCAGTGCCCTGTGATCATT
TCAAGGGCAATGTGAAGAAAACAAGACACCAAAGGCACCACAGAAAGCCAACAAAGCATTCCAGA
GCCTGCCAGCAATTCTCAAACAATGTCAGCTAAGAAGCTTGCTCTGCCTTGTAGGAGCTCTG
AGCGCCCACCTTCCAATTAAACATTCTCAGCCAAGAAGACAGTGAGCACACCTACCAGACACTC
TTCTTCTCCACCTCACTCTCCACTGTACCCACCCCTAAATCATCCAGTGCTCTCAAAAGCA
TGTTTTCAAGATCATTGTTGCTCTCTAGTGTCTTCTCTCGTCAGTCTTAGCCT
GTGCCCTCCCTTACCCAGGCTAGGCTTAATTACCTGAAAGATTCCAGGAAACTGTAGCTTCC
AGCTAGTGTCAATTAAACCTTAAATGCAATCAGGAAAGTAGCAAACAGAAGTCAATAAATATTTT
AAATGTCAAAAAAAAAAAAAAAA

FIGURE 26

MKVLISLLLLPLMLMSMVSSSLNPGVARGHRDRGQASRRWLQEGGQECECKDWFLRAPRRKFM
TVSGLPKKKQCPDCDFKGNVKKTRHQRHHRKPNKHSRACQQFLKQCQLRSFALPL

Important features:

Signal peptide:

amino acids 1-22

N-myristoylation sites.

amino acids 27-33, 46-52

FIGURE 27

GGACGCCAGGCCCTGCAGAGGCTGAGCAGGGAAAAAGCCAGTCCCCAGCGGAAGCACAGCTCAG
AGCTGGTCTGCCATGGACATCCTGGTCCACTCCTGCAGCTGCTGGTCTGCTTACCTGCC
CCTGCACCTCATGGCTCTGCTGGCAGCCCCTGTGAAAAGCTACTTCCCTACCTGA
TGGCCGTGCTGACTCCAAAGAGCAACCGCAAGATGGAGAGCAAGAACGGGAGCTTCAGCCAG
ATAAAAGGGCTTACAGGAGCCTCCGGAAAGTGGCCCTACTGGAGCTGGCTGCCAACCGGAGC
CAACTTCAGTTCTACCCACCGGGCTGCAGGGCACCTGCCTAGACCCAAATCCCCACTTGAGA
AGTTCTGACAAAGAGCATGGCTGAGAACAGGCACCTCCAATATGAGCGGTTGTGGCTCCT
GGAGAGGACATGAGACAGCTGGCTGATGGCTCCATGGATGTGGTCTGCACTCTGGTGTG
CTCTGTGCAGAGCCAAGGAAGGTCTGCAGGAGGTCCGGAGAGTACTGAGACCGGGAGGTGTG
TCTTTCTGGGAGCATGTGGCAGAACCATATGGAAGCTGGCCTCATGTGGCAGCAAGTTTC
GAGCCCACCTGGAAACACATTGGGATGGCTGCTGCCCTACCAAGAGAGACCTGGAAGGATCTGA
GAACGCCAGTTCTCCAAATGGAACGACAGCCCCCTCCCTGAAGTGGTACCTGTTG
GGCCCCACATCATGGAAAGGCTGTCAAACAAATCTTCCAAGCTCCAAGGCACACTTGTCC
TTCCCCAGCCTCCAATTAGAACAGCCACCCACCAGCCTATCTATCTTCACTGAGAGGGACTA
GCAGAACATGAGAGAAGACATTCATGTACCACTACTAGTCCTCTCTCCCCAACCTCTGCCAGGGC
AATCTCTAACTCAATCCGCCTCGACAGTGAAGGGCTACTCTACGCTGACCCAGGGAGG
AAACACTAGGACCCCTGTTGTATCCTCAACTGCAAGTTCTGGACTAGTCTCCCAACGTTGCC
CCAATGTTGCCCTTCCTCGTTCCATGGTAAAGCTCCTCTCGCTTCCCTGAGGCTACAC
CCATGCGTCTCTAGGAACGGTCACAAAGTCATGGTGCTGCATCCCTGCCAACCCCCCTGAC
CCTCTCTCCCCACTACCACCTCTTGAGCTGGGGGACCAAGGGAGAATCAGAGATGCTGGGG
ATGCCAGAGCAAGACTCAAAGAGGCAGAGGTTGTTCTCAAATATTTTAATAAGACGAA
ACCACG

FIGURE 28

MDILVPLLQLLVLLTLPLHLMALLGCWQPLCKSYFPYLMAVLTPKSNRKMESKKRELFSQIKGL
TGASGKVALLELGCCTGANFQFYPPGCRVTCLDPNPHFEKFLTKSMAENRHLQYERFVVAPGEDM
RQLADGSMDVVVCTLVLCVQSPRKVLQEVRVLRPGGVLFWEHVAEPYGSWAFMWQQVFEPTW
KHIGDGCCLTRETWKDLENAQFSEIQMERQPPPLKWLPGPHIMGKAVKQSFPSKALICSFPSL
QLEQATHQPIYLPLRGT

Important features:

Signal peptide:

amino acids 1-23

Leucine zipper pattern.

amino acids 10-32

N-myristoylation sites.

amino acids 64-70, 78-84, 80-86, 91-97, 201-207

FIGURE 29

CAATGTTGCCATCCACCTCCCCAAGCCCCTTACCTATGCTGCTGCTAACGCTGCTGCTGCT
GCTGCTGCTGCTAAAGGCTCATGCTTGGAGTGGGACTGGTCGGGCCAGAAAGTCTCTTCTG
CCACTGACCCCCCATCAGGGATTGGGCCTTCTTCCCCCTCCTTCTGTGTCTCCTGCCTCAT
CGGCCTGCCATGACCTGCAGCCAAGCCCAGCCCCGTGGGAAGGGAGAAAGTGGGGATGGCTA
AGAAAGCTGGGAGATAGGGAACAGAACAGAGGGTAGTGGGTGGCTAGGGGGCTGCCTATTTAAA
GTGGTTTTATGATTCTTACTAATTATAACAAAGATATTAAGGCCCTGTTCAATTAAAGAAATT
GTTCCCTCCCTGTGTTCAATGTTGAAAGATTGTTCTGTGTAATATGTCTTATAATAAC
AGTTAAAAGCTGAAAAAAAAAAAAAAAAAAAAAA

FIGURE 30

MLLLTLLLLLKGSCLEGLVGAQKVSSATDAPIRDWAFFPPSFLCLLPHRPAMTCSQAQPRG
EGEKVGDG

Important features:

Signal peptide:

amino acids 1-15

Growth factor and cytokines receptors family:

amino acids 3-18

FIGURE 31

GTTCGAATTCCCTCAACTATAACCCACAGTCCAAAAGCAGACTCACTGTGTCCCAGGCTACCAGTT
CCTCCAAGCAAGTCATTCCTTATTAACCGATGTGTCCCTCAAACACCTGAGTGCTACTCCCT
ATTTGCATCTGTTGATAAATGATGTTGACACCCCTCCACCGAATTCTAAGTGAATCATGTCGG
GAAGAGATAACAATCCTGGCCTGTGTATCCTCGCATTAGCCTGTCTTGGCCATGATGTTACC
TTCAGATTCATCACCAACCCCTCTGGTTCACATTTCATTGTTATTGGGATTGTTGTT
TGTCTGCGGTGTTATGGTGGCTGTATTATGACTATAACCAACGACCTCAGCATAGAATTGGACA
CAGAAAGGAAAATATGAAGTGCCTGCTGGGTTGCTATCGTATCCACAGGCATCACGGCAGTG
CTGCTCGTCTGATTTGTTCTCAGAAAGAGAATAAAATTGACAGTTGAGCTTCCAAATCAC
AAATAAGCCATCAGCAGTGCTCCCTCCTGCTGTTCCAGCCACTGTGGACATTGCCATCCTCA
TTTCTTCTGGTCCTCTGGTGGCTGTGCTGAGCCTGGAACTGCAGGAGCTGCCAGGTT
ATGGAAGGCGGCCAAGTGAATATAAGCCCCTTCGGGCATTGGTACATGTGGTGTACCATTT
AATTGGCCTCATCTGGACTAGTGAATTCATCCTTGCCTGCCAGCAAATGACTATAGCTGGGCAG
TGGTACTTGTATTCAACAGAAGTAAAATGATCCTCCTGATCATCCCATTCTCGTCTCTC
TCCATTCTCTTCTACCATCAAGGAACCGTTGAAAGGGTATTAACTCTGTGGTGTAG
GATTCCGAGAATCATTGTCATGTACATGCAAACGCACTGAAAGAACAGCAGCATGGTCATTGT
CCAGGTACCTGTTCCGATGCTGACTGCTGTTCTGGTGTCTGACAAATACCTGCTCCATCTC
AACCCAGAATGCATATACTACAACGTCTATTAAATGGGACAGATTCTGTACATCAGCAAAGATGC
ATTCAAATCTTGTCCAAGAACTCAAGTCACTTACATCTATTAACTGCTTGGAGACTTCATAA
TTTTCTAGGAAAGGTGTTAGGGTGTGTTCACTGTTGGAGGACTCATGGCTTTAACTAC
AATCGGGCATTCCAGGTGTGGCAGTCCCTCTGTTATTGGTAGCTTTGCCTACTTAGTAC
CCATAGTTTTATCTGTGTTGAAACTGTGCTGGATGCACCTTCTGTGTTGCTGTTGATC
TGGAAACAAATGATGGATCGTCAGAAAAGCCCTACTTATGGATCAAGAATTCTGAGTTGTA
AAAAGGAGCAACAAATTAAACAATGCAAGGGCACAGCAGGACAAGCACTCATTAAGGAATGAGGA
GGGAACAGAACTCCAGGCCATTGTGAGATAGATACCCATTAGGTATCTGTACCTGGAAAACATT
TCCTTCTAAGAGCCATTACAGAATAGAAGATGAGACCACTAGAGAAAAGTTAGTGAATT
TTAAAAGACCTAATAAACCTATTCTCCTCAAAA

FIGURE 32

MSGRDTI~~L~~GLCILALALSLAMMFTFRFITTLLVHIFISLVLG~~L~~FVCGVLWWLYDYTN~~D~~L~~S~~IE
LDTERENMKC~~V~~LGFAIVSTGITAVLLVLIFVLRKRIKLTVELFQITNKA~~I~~SSAPF~~L~~FQPLWTFA
ILIFFWVLWAV~~V~~AVLLSLGTAGAAQVMEGGQ~~V~~EYKPLSGIRYMWSYH~~L~~IGLIWTSEF~~I~~LA~~C~~Q~~Q~~MTIA
GAVVTCYFNR~~S~~KNDPPDHP~~I~~LSSLSILFFYHQGT~~V~~VKG~~S~~FLISV~~V~~RIPRIIVM~~Y~~M~~Q~~NALKEQQHG
ALSRYLF~~R~~C~~C~~YCCFWCLDKYLLH~~L~~NQNAYTTAINGTD~~F~~CTS~~A~~DAFK~~I~~LSKN~~S~~HF~~T~~SINC~~F~~GD
FII~~I~~FLGKVLVV~~C~~FTVFG~~G~~LM~~A~~F~~N~~YNRAFQV~~W~~AV~~P~~LLVAFFAYLVAHSFLSV~~F~~ETVLDALFLCFA
VDLETNDGSSEKPYFMDQEFLSFVKRSNKLNNARAQQDKHSLRNEEGTELQAIVR

Important features:

Signal peptide:

amino acids 1-20

Putative transmembrane domains:

amino acids 35-54, 75-97, 126-146, 185-204, 333-350, 352-371

N-glycosylation sites.

amino acids 204-208, 295-299, 313-317

N-myristoylation sites.

amino acids 147-153, 178-184, 196-202, 296-275, 342-348

FIGURE 33

GTTCGATTAGCTCCTCTGAGAAGAAGAAAAGTTCTGGACCTCCCTGTTCTCCTAGA
ATAATTGTATGGATTGTGATGCAGGAAAGCTAAGGAAAAAGAATATTCTGTGTGGT
GAAAATTTGAAAAAAAATTGCCTCTTCAAACAAGGGTGTATTCTGATATTATGAGGAC
TGTTGTTCTCACTATGAAGGCATCTGTATTGAAATGTTCTTGTGACTGGAGTAC
ATTCAAACAAAGAACGGCAAAGAAGATTAAAAGGCCAAGTTCAGTGTGCCTCAGATCAACTGC
GATGTCAAAGCCGAAAGATCATCGATCCTGAGTTCATTGTGAAATGTCCAGCAGGATGCCAAGA
CCCCAAATACCATGTTATGGCACTGACGTGTATGCATCCTACTCCAGTGTGTGGCGCTGCCG
TACACAGTGGTGTGCTTGATAATTAGGAGGGAAAATACTTGTTCGGAAGGGTGCTGGACAGTCT
GGTACAAAGGGAGTTATTCCAACGGTGTCCAATCGTTATCCCTACCACGATGGAGAGAATCCTT
TATCGTCTAGAAAGTAAACCCAAAAGGGTGTAAACCTACCCATCAGCTTACATACTCATCAT
CGAAAAGTCCAGCTGCCAAGCAGGTGAGACCACAAAAGCCTATCAGAGGCCACCTATTCCAGGG
ACAAC TG CACAGCCGGTCACTCTGATGCAGCTCTGGCTGTCACTGTAGCTGTGGCCACCCCCAC
CACCTTGCCAAGGCCATCCCCTCTGCTGCTTCTACCACCAAGCAGGCCACAAATCAGTGG
GCCACAGGAGCCAGGAGATGGATCTCTGGTCCACTGCCACCTACACAAGCAGGCCAAACAGGCC
AGAGCTGATCCAGGTATCCAAAGGAAGATCCTTCAGGAGCTGCCTCCAGAAACCTGTTGGAGC
GGATGT CAGCCTGGGACTTGTCCAAAAGAAGAATTGAGCACACAGTCTTGGAGGCCAGTATCCC
TGGGAGATCCAAACTGCAAATTGACTTGTGTTTAATTGATGGGAGCACCAGCATTGGCAA
CGGCGATTCCAATCCAGAAGCAGCTCTGGCTGATGTTGCCAAGCTTGCACATTGGCCCTGC
CGGTCCACTGATGGGTGTTGTCAGTATGGAGACAACCCCTGCTACTCACTTAACCTCAAGACAC
ACACGAATTCTCGAGATCTGAAGACAGCCATAGAGAAAATTACTCAGAGAGGGAGCTTCTAAT
GTAGGTGGGCCATCTCCTTGTGACCAAGAACTTCTTCCAAAGCCAATGGAAACAGAAGCGG
GGCTCCAATGTGGTGGTGTGATGGTGGATGGCTGGCCACGGACAAAGTGGAGGGCTTCAA
GACTTGCAGAGAGTCAGGAATCAACATTCTCATCACCATTGAAGGTGCTGCTGAAATGAG
AAGCAGTATGTGGTGGAGGCCAACTTGCAAACAAGGCCGTGTGCAGAACAAACGGCTTCTACTC
GCTCCACGTGCAGAGCTGGTTGGCCTCCACAAGACCCCTGCAGCCTCTGGTGAAGCAGGGCTGCG
ACACTGACCGCCTGGCCTGCAGCAAGACCTGCTGAACCTGGCTGACATTGGCTCGTACGAC
GGCTCCAGCAGTGTGGGACGGCAACTCCGACCGTCCAGTTGTGACCAACCTCACCAA
AGAGTTGAGATTCCGACACGGACACGCGCATGGCCGTGCAGTACACCTACGAACAGCAGGC
TGGAGTTGGGTCAGCAAGTACAGCAGCAAGCCTGACATCCTCAACGCCATCAAGAGGGTGGC
TACTGGAGTGGTGGCACCAGCACGGGGCTGCCATCAACTCGCCCTGGAGCAGCTTCAAGAA
GTCCAAGCCAAACAAGAGGAAGTTAATGATCCTCATCACCAGGGAGGTCTACGACGACGTCC
GGATCCCAGCCATGGCTGCCATCTGAAGGGAGTGATCACCTATGCAGTAGGCAGTGGCTGGCT
GCCCAAGAGGAGCTAGAAGTCATTGCCACTCACCCGCCAGAGACCAACTCCTCTTGTGGAGCA
GTTTGACAACCTCCATCAGTATGTCCCCAGGATCATCCAGAACATTGTACAGAGTTCAACTCAC
AGCCTCGGAACTGAATTCAAGAGCAGGCAGAGCACCAGCAAGTGTGCTTACTAAGTGTGTT
GGACCAACCCACCGCTTAATGGGGCACGCACGGTGCATCAAGTCTGGCAGGGCATGGAGAAC
AAATGTCTTGTATTATTCTTGCCATCATGCTTTCATATTCCAAAATGGAGTTACAAAGA
TGATCACAAACGTATAAGAATGAGCAAAGGCTACATCATGTTGAGGGTGTGGAGATTTCAT
TTTGACAATTGTTCAAATAAATGTCGAATACAGTGCAGCCCTACGACAGGCTTACGTAG
AGCTTTGTGAGATTAAAGTGTATTGCTGAAATTAAATGTTAGAAGGATGAAAATAAAAAAA
AAA
AAAAAAAAAAAAAAAAAAAAAAAAAAG

FIGURE 34

MRTVVLTMKASVIEMFLVLLVTGVHSNKETAKKIKRPKFTVPQINC DVKAGKIIDPEFIVKCPAG
CQDPKYH VY GTDVY ASYSSVCGAAVHSGVLDNSGGKILVRKVAGQSGYKGSYSNGVQSLSLPRWR
ESFIVLESKKGVTVPSALTYSSSKSPAAQAGETTKAYQRPPIPGTTAQPVTLMQLLAVTVAVA
TPPTLPRPSPSAASTTSIPRPQSVGHRSQEMDLWSTATYTSSQNPRADPGIQRQDPSGAAFQKP
VGADVSLGLVPKEELSTSLEPVSLGDPNCKIDLSFLIDSTSIGKRRFRIQKQLLADVAQALDI
GPAGPLMGVVQYGDNPATHFNLKHTNSRDLKTAIEKITQRGGLSNVGRAISFVTKNFFSKANGN
RSGAPNVVVVMVDGWPTDKVEEASRLARESGINIFFITIEGAAENEKQYVVEPNFANKAVCRTNG
FYSLHVQSWFGLHKTQPLVKRVCDTDLACSKTCLNSADIGFVIDGSSVGTGNFRTVLQFVTN
LTKEFEISDTDTRIGAVQYTYEQRLEFGFDKYSSKPDILNAIKR VGYWSGGTSTGAAINFALEQL
FKKS KPNKRKLMILITDGRSYDDVRIPAMA AHLKGVITYAIGVAWAQEELEVIATHPARDHSFF
VDEF DNLHQYVPRIIQNICTEFNSQPRN

Important features:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 181-200

N-glycosylation sites.

amino acids 390-394, 520-524

N-myristylation sites.

amino acids 23-29, 93-99, 115-121, 262-268, 367-373, 389-395,
431-437, 466-472, 509-515, 570-576, 571-577, 575-581, 627-633

Amidation site.

amino acids 304-308

FIGURE 35

CCGAGCACAGGAGATTGCCTGCGTTAGGAGGTGGCTGCGTTGGAAAAGCTATCAAGGAAGAAATTGC
CAAACCATGTCTTTCTGTTTCAGAGTAGTTACAACAGATCTGAGTGTAAATTAAGCATGGAAT
ACAGAAAACAACAAAAACTTAAGCTTAATTCATCTGGAATCCACAGTTCTAGCTCCCTGGACCC
GGTGACCTGTTGGCTCTCCCGCTGGCTGCTCTATCACGTGGTGCCTCCGACTACTCACCCGAGTGT
AAGAACCTCGGCTCGCGTGCCTGAGCTGCTGGATGGCCTCGGCTCTGGACTGTCCTCCGAGTA
GGATGTCACTGAGATCCCTCAAATGGAGCCTCCTGCTGCTGTCACTCCTGAGTTCTTGATGTGGTAC
CTCAGCCTCCCCACTACAATGTGATAGAACCGGTGAACGGATGTACTCTATGAGTATGAGCCGATT
CAGACAAGACTTTCACTTCACACTCGAGAGCATTCAAACGCTCTCATAAAATCCATTCTGGTCATTC
TGGTGACCTCCCACCCCTCAGATGTGAAAGGCCAGGCCATTAGAGTTACTTGGGTGAAAAAAAGTCT
TGGTGGGATATGAGGTTCTTACATTTCTTATTAGGCCAAGAGGCTGAAAGGAAGACAAAATGTTGGC
ATTGTCCTTAGAGGATGAACACCTCTTATGGTACATAATCCGACAAGATTTTAGACACATATAATA
ACCTGACCTTGAAAACCATTATGCCATTCAAGGTGGTAACTGAGTTGCCCAATGCCAAGTACGTAATG
AAGACAGACACTGATGTTCAACTGGCAATTAGTGAAGTATCTTAAACCTAAACCACTCAGA
GAAGTTTCACAGGTATCCTCTAATTGATAATTATTCTATAGAGGTTTACCAAAAAACCCATATT
CTTACCAAGGAGTATCCTCAAGGTGTTCCACTGCAGTGGTTATATAATGCCAGAGAT
TTGGTGCCAAAGGATCTATGAAATGATGGTCACGTAAAACCCATCAAGTTGAAGATGTTATGTCGGGAT
CTGTTGAATTATTAAAAGTGAACATTCAATTCCAGAACAGACACAAATCTTCTTCTATAGAATCC
ATTGGATGTCGTAACGTGAGACGTGTGATTGCAGCCATGGCTTTCTTCCAAGGAGATCATCACTTT
TGGCAGGTCAAGAACACCACATGCCATTATTAACTTCACATTCTACAAAAGCCTAGAAGGACAG
GATACCTTGTGAAAGTGTAAATAAAAGTAGGTACTGTGAAAATTCAATGGGAGGTCAGTGTGCTGGCTT
ACACTGAACTGAAACTCATGAAAAACCCAGACTGGAGACTGGAGGGTTACACTTGTGATTTATTAGTCAGG
CCCTCAAAGATGATATGTGGAGGAATTAAATATAAGGAATTGGAGGTTTGTCAAAGAAATTAAATAGG
ACCAACAAATTGGACATGTCATTCTGTAGACTAGAATTCTAAAGGGTGTACTGAGTTAAAGCTCA
CTAGGCTGAAAAACAAACATGTAGAGTTTATTGAAACAATGTAGTCACGTGAAGGTTGT
TATCTTATGTGGATTACCAATTAAAAATATGTAGTTCTGTCAAACAAACTTCTTCACTGAAGTTATA
CTGAACAAAATTTCACCTGTTGGTCATTATAAGTACTTCAGATGTTGCAGTATTCACAGTTATT
ATTATTAAAATTACTCAACTTGTGTTTAAATGTTGACGATTCAATACAAGATAAAAGGATAG
TGAATCATTCTTACATGCAAACATTCCAGTTACTTAACGTGATCAGTTATTGATAACATCACTCCA
TTAATGTAAAGTCATAGGTCAATTGCAATCAGTAATCTGGACTTGTAAATATTACTGTGGT
AATATAGAGAAGAATTAAAGCAAGAAAATCTGAAAA

FIGURE 36

MASALWTVLPSRMSLRSLKWSLLLLSLLSFFVMWYLSLPHYNVIERVNWMYFYEYEPIYRQDFHF
TLREHSNC SHQNPFLVILVTSHPSDVKARQAIRVTWGEKKSWWGYEVLTFFLLGQEAEKEDKMLA
LSLEDEHLLYGDII RQDFLDTYNNLT LKTI MAFRWVTEFCPNAKYVMKTDTDVFINTGNLVKYLL
NLNHSEKFFTGYPLIDNYSYRGFYQKTHISYQEYPKVFPPYCSGLGYIMSRLVPRIYEMMGHV
KPIKFEDVYVGICLNLLKVNIHIPEDTNLFFLYRIHLDVCQLRRVIAAHGFSSKEIITFWQVMLR
NTTCHY

Important features:

Type II transmembrane domain:

amino acids 20-39

N-glycosylation sites.

amino acids 72-76, 154-158, 198-202, 212-216, 326-330

Glycosaminoglycan attachment site.

amino acids 239-243

Ly-6 / u-PAR domain proteins.

amino acids 23-37

N-myristoylation site.

amino acids 271-277

FIGURE 37

CGCTCGGGCACCAAGCGGGCAAGGATGGAGCTGGGTTGCTGGACGCAGTGGGGCTCACTTTCTTCAGCTCCTCTCATC
TCGTCTTGCAAGAGAGTACACAGTCATTAATGAAGCCTGCCCTGGAGCAGAGTGGAAATATCATGTGTCGGAGTGCTGTG
AATATGATCAGATTGAGTGCGTCTGCCCGGAAAGAGGGAAAGTCGTGGTTATACCATCCCTGCTGCAGGAATGAGGAGAA
TGAGTGTGACTCCTGCCTGATCCACCCAGGTTGACCATCTTGAAAAGTCAAGAGCTGCCAATGGCTATGGGGGT
ACCTTGGATGACTTCTATGTGAAGGGTTCTACTGTGCAGAGTGCGAGCAGGCTGGTACGGAGGAGACTGCATGCGATGTG
GCCAGGTTCTGCCAGCCCCAAAGGGTCAGATTTGGAAAGCTATCCCTAAATGCTACTGTGAATGGACCATTGC
TAAACCTGGGTTGTCATCCAACTAAGATTGTCATGTTGAGTCAGGACTACATGTGCCAGTATGACTATGTTGAG
GTTCGTGATGGAGACAACCGCGATGGCAGATCATCAAGCGTGTCTGAGCAACGAGCGGCCAGCTCTATCCAGAGCAG
GATCCTCACTCCACGTCTCCACTCCGATGGCTCAAGAATTGACGGTTCCATGCCATTATGAGGAGATCACAGC
ATGCTCCTCATCCCTGTTCCATGACGGCACGTGCGTCTGACAAGGCTGGATCTTACAAGTGTGCCTGCTGGCAGGC
TATACTGGGCAGCGTGTGAAAATCTCCTGAAAGAAAAGACTGCTCAGACCCCTGGGGCCAGTCAGTGGTACAG
TAACAGGGGCCCTGGCTTATCAACGGACGCCATGCTAAAATTGGCACCGTGGTCTTCTTTGTAACAACCTCTATGT
TCTTAGTGGCAATGAGAAAAGAAACTGCCAGCAGAATGGAGAGTCTCCGATGCAGGTTCAAGGGAGACACCATTACACCAGCTAT
GAACCAAAGATTCAACGACTGGTGAGAAGGAGAGTCTCCGATGCAGGTTCAAGGGAGACACCATTACACCAGCTAT
ACTCAGCGGCCTTCAGCAAGCAGAAAAGTGCAGAGTGCCTTACCAAGAAGCCAGGCCCTCCCTGGAGATCTGCCATGGG
ATACCAACATCTGCATACCCAGCTCCAGTATGAGTCATCTCACCCCTTACCGCCGCTGGCAGCAGCAGGAGGACATGT
CTGAGGACTGGGAAGTGGAGTGGCGGGCACCACCTGCATCCCTATCTGGGAAAATTGAGAACATCACTGCTCCAAAGA
CCCAAGGGTTGCCTGGCGTGGCAGGCCATCTACAGGAGGACCAGCAGGGTGCACTGACGGCAGCCTACACAAGGGAGC
GTGGTCTCTAGTCTGCAGCGGTGCCCTGGTAATGAGCCACTGTGGTGGCTGCCACTGTGTTACTGACCTGGGAAG
GTCACCATGATCAAGACAGCAGACCTGAAAGTTGGGAAATTCTACCGGGATGATGACGGGATGAGAACGACCATCC
AGAGCCTACAGATTCTGCTATCATTCTGCATCCAACTATGACCCATCCTGCTGATGCTGACATGCCATCCTGAAGCT
CCTAGACAAGGCCGTATCAGCACCCGAGTCCAGCCCCTCGCTGCCAGTCAGGAGCAGCAGGAGGACATGCCCTGGG
TCCCACATCACTGTGGCTGGCTGGAATGTCCTGGCAGACGTGAGGAGGCCCTGGCTCAAGAACGACACACTGCGCTCTGGG
TGGTCAGTGTGGGACTCGCTGCTGTGAGGAGCAGCATGAGGAGCAGTCAGTGTGACTGATAACATGTT
CTGTGCCAGCTGGGAACCCACTGCCCTCTGATATCTGCAGAGACAGGAGGACATCGCGCTGTGCTCTCCGGGA
CGAGCATCTCTGAGCCACGCTGGCAGTGTGGACTGGTCAGCTGGAGCTATGATAAAACATGCCACAGGCTCTCCA
CTGCCCTCACCAAGGTGCTGCCCTTAAAGACTGGATTGAAAGAAAATGAAACCATGCTCATGCACTCCTGAGAAG
TGTTTCTGTATATCCGTCTGTACGTGTGTCATTGCGTGAAGCAGTGTGGGCTGAAGTGTGATTTGGCCTGTGAACTGGCT
GTGCCAGGGCTCTGACTCAGGGACAAAACCTCAGTGAAGGGTAGACCTCCATTGCTGGTAGGCTGATGCCCGTCCA
CTACTAGGACAGCCAATTGGAAGATGCCAGGGCTTGCAGAGAAGTAAGTTCTCAAAGAACCCATATACAAACCTCTCCA
CTCCACTGACCTGGTGGCTTCCCAACTTCAGTTACGAATGCCATCAGCTTGACCCAGGGAGATCTGGCTTCTGAG
GCCCTTGGAGGCTCTCAAGTTCTAGAGAGCTGCCGTGGGACAGGCCAGGGCAGCAGAGCTGGGATGTGGTGCATGCC
TGTGTACATGGCACAGTACAGTCTGGCTTCCCTCCATCTTGTACACATTAAATAAAAATAAGGGTTGGCTCT
GAACTACAA
AA

FIGURE 38

MELGCWTQLGLTFLQLLLISLPREYTGINEACPGAEWNIMCRECCEYDQIECVCPGKREVVGYT
IPCCRNEENECDSCLIHPGCTIFENCKSCRNGSWGGTLDDFYVKGFYCAECRAGWYGGDCMRCQ
VLRAPKGQILLESYPLNAHCEWTIHAKPGFVIQLRFVMLSLEFDYMCQYDYVEVRDGDNRDGQII
KRVCGNERPAPIQSIGSSLHVLFHSDGSKNFDGFHAIYEEITACSSSPCFHDGTCVLDKAGSYKC
ACLAGYTGQRCENLLEERNCSDPGGPVNGYQKITGGPGLINGRHAKIGTVVSFFCNNSYVLSGNE
KRTCQQNGEWSGKQPIKACREP KISDLVRRVLPMQSRETPLHQLYSAAFSKQKLQSAPTK
KPALPFGDLPMGYQHLHTQLQYECISPFYRRLGSSRRTCLRTGKWSGRAPSCIPICGKIE NITAP
KTQGLRWPWQAAIYRRTSGVHDGSLHKGAWFLVCSGALVNERTVVVAHCVTDLGKVTMIKTADL
KVVLGKFYRDDDREKTIQSLQISAIILHPNYDPILLDADIAILKLLDKARI STRVQPI CLAASR
DLSTS FQESHITVAGWNVLADVRSPGFKN DTLRSGVVSVVDSLLCEEQHEDHGIPVSVTDNMFCA
SWEPTAPS DICTAETGGIAAVSFPGRASPEPRWHL MGLVWSYDKTC SHRLSTAFTKVL PFKDWI
ERNMK

Important features of the protein:

Signal peptide:

amino acids 1-23

EGF-like domain cysteine pattern signature.

amino acids 260-272

N-glycosylation sites.

amino acids 96-100, 279-283, 316-320, 451-455, 614-618

N-myristoylation sites.

amino acids 35-41, 97-103, 256-262, 284-290, 298-304, 308-314,
474-480, 491-497, 638-644, 666-672

Amidation site.

amino acids 56-60

Serine proteases, trypsin family.

amino acids 489-506

CUB domain proteins profile.

amino acids 150-167

FIGURE 39

FIGURE 40

MKALLLVLPWLSPANYIDNVGNLHFLYSELCKGASHYGLTKDRKRRSQDGCPDGCA
SLTATAPS
PEVSAATISLMTDEPGLDNPAYVSSAEDGQPAISPVDGSRNRTRARPFERSTIRSR
SFKKINR
ALSVLRRTKSGSAVANHADQGRENSENTTAPEVFPRLYHILPDGEITSIKINRVDP
SESLSIRLV
GGSETPLVHIIQHIYRDGVIARDGRLLPGDIILKVNGMDISNVPHNYAVRLLRQPC
QVLWLTVM
REQKFRSRNNGQAPDAYRPRDDSFHVLINKSSPEEQLGIKLVRKVDEPGVFI
FNVLDDGVAYRHG
QLEENDRVLAINGHDLRYGSPESAHLIQASERRVHLVVSQRQSPDI
FQEAGWNSNGSWSPG
PGERSNTPKPLHPTITCHEKVVNIQKDPGESLGMTVAGGASHREWDLPIYVIS
VEPGVISRDGR
IKTGDILLNVDGVELTEVRSEAVALLKRTSSSIVLKALEVKEYEPQEDCSSPA
ALDSNHNMAPP
SDWSPSWVMWLELPRCLYNCKDIVLRRNTAGSLGFCIVGGYEEYNGNKPFFIKS
IVEGTPAYNDG
RIRCGDILLAVNGRSTSGMIHACLRLLKELKGRITLTIVSWPGTFL

Important features:

Signal peptide:

amino acids 1-15

N-glycosylation sites.

amino acids 108-112, 157-161, 289-293, 384-388

Tyrosine kinase phosphorylation sites.

amino acids 433-441, 492-500

N-myristoylation sites.

amino acids 51-57, 141-147, 233-239, 344-350, 423-429, 447-453,
467-473, 603-609

FIGURE 41

ACCAGGCATTGTATCTTCAGTTGTCAAGTCGCAATCAGATTGGAAAAGCTCAACTGAAGCTT
CTTGCCTGCAGTGAAGCAGAGAGATAGATATTATTACGTAATAAAAACATGGGCTTCAACCTGACT
TTCCACCTTCTACAAATTCCGATTACTGTTGCTGTTGACTTGTGCCTGACAGTGGTTGGGTGGC
CACCAAGTAACACTTCGTGGGTGCCATTCAAGAGATTCCCTAAAGCAAAGGAGTCATGGCTAATTCC
ATAAGACCCCTCATTGGGGAAAGGGAAAAACTCTGACTAATGAAGCATTCCACGAAGAAGGTAGAACTT
GACAACGTCCCTCTGTCCTACCTCAGAGGCCAGAGCAAGCTCATTCAAACCCAGATCTCAC
TTTGGAAAGAGGTACAGGCAGAAAATCCCAAAGTGTCCAGAGGCCGGTATGCCCTCAGGAATGTAAG
CTTACAGAGGGTCGCCATCCTCGTCCCCACCGGAACAGAGAGAAACACCTGATGTACCTGCTGGAA
CATCTGCATCCCTCCTGCAGAGGCAGCAGCTGGATTATGGCATCTACGTCACTCACCAGGCTGAAGG
TAAAAAGTTAATCGAGCCAAACTCTGAATGTGGCTATCTAGAACCCCTCAAGGAAGAAAATTGGG
ACTGCTTATATTCCACGATGTGGACCTGGTACCCGAGAATGACTTTAACCTTACAAGTGTGAGGAG
CATCCCAAGCATCTGGTGGTGGCAGGAACAGCACTGGTACAGGTTACAGTGGATATTGG
GGGTGTTACTGCCCTAACAGAGAGCAGTTTCAAGGTGAATGGATTCTCTAACAAACTACTGGGAT
GGGGAGGCGAAGACGATGACCTCAGACTCAGGGTTGAGCTCCAAAGAATGAAAATTCCCGCCCTG
CCTGAAGTGGTAAATATAACATGGCTTCCACACTAGAGACAAAGGAATGAGGTGAACGCAGAACG
GATGAAGCTTACACCAAGTGTACGAGTCTGGAGAACAGATGGTTGAGTAGTTGTTCTATAAAT
TAGTATCTGTGGAACACAATCCTTATATATCAACATCACAGTGGATTCTGGTTGGTGCATGACCC
TGGATCTTGGTGTGTTGGAAGAACTGATTCTTGTGCAATAATTGGCCTAGAGACTTCAA
ATAGTAGCACACATTAAGAACCTGTTACAGCTATTGTTGAGCTGAATTTCCTTTGTATTTCT
TAGCAGAGCTCCTGGTGTAGAGTATAAACAGTTGTAACAAAGACAGCTTCTTAGTCATTTGAT
CATGAGGGTTAAATATTGTAATATGGATACTTGAAGGACTTATATAAAAGGATGACTCAAAGGATAA
AATGAACGCTATTGAGGACTCTGGTTGAAGGAGATTATTAAATTGAAGTAATATATTGGGAT
AAAAGGCCACAGGAAATAAGACTGCTGAATGTCAGAGAGAACCAAGAGTTGTTCTCGTCCAAGGTAGAA
AGGTACGAAGATAACAATACTGTTATTCAATTCTGTACAATCATCTGTGAAGTGGTGGTGTCAAGGT
GAGAAGGCGTCCACAAAGAGGGAGAAAGGCAGCAATCAGGACACAGTGAACCTGGGAATGAAGA
GGTAGCAGGAGGGTGGAGTGTGGCTGCAAAGGCAGCAGTAGCTGAGCTGGTTGCAGGTGCTGATAGC
CTTCAGGGAGGACCTGCCAGGTATGCCCTCAGTGTGATGCCACCAGAGAAATACATTCTATTAGT
TTTAAAGAGTTTGTAAAATGATTGTACAAGTAGGATATGAATTAGCAGTTACAAGTTACAT
ATTAACATAATAATGTCTATCAAATACCTCTGTAGTAAAATGTGAAAAGCAAA

FIGURE 42

MGFNLTFHLSYKFRLLLLLTLCLTVVGWATSNYFVGAIQEIPKAKEFMANFHKTLLGKGKTLTN
EASTKKVELDNCPSVSPYLRGQSKLIFKPDLTLEEVQAENPKVSRGRYRPQECKALQRVAILVPH
RNREKHLMYLLEHLHPFLQRQOLDYGIYVIHQAEGKKFNRAKLLNVGYLEALKEENWDCFIFHDV
DLVPENDFNLYKCEEHPKHLVVGRNSTGYRLRYSGYFGGVTAALSREQFFKVNNGFSNNYWGWGGED
DDLRLRVELQRMKISRPLPEVGKYTMVFHTRDKGNEVNAERMKLLHQVSRVWRTDGLSSCSYKLV
SVEHNPLYINITVDFWFGA

Important features:

Signal peptide:

amino acids 1-27

N-glycosylation sites.

amino acids 4-8, 220-224, 335-339

Xylose isomerase proteins.

amino acids 191-202

FIGURE 43

GCTCAAGACCCAGCAGTGGGACAGCCAGACAGACGGCACGATGGCACTGAGCTCCCAGATCTGGG
CCGCTTGCCTCCTGCTCCTCCTCGCCAGCCTGACCAGTGGCTCTGTTTCCCACAACAG
ACGGGACAACCTGCAGAGCTGCAACCCCAGGACAGAGCTGGAGCCAGGGCCAGCTGGATGCCCAT
GTTCCAGAGGCGAAGGAGGCGAGACACCCACTTCCCCATCTGCATTTCTGCTGCGGCTGCTGTC
ATCGATCAAAGTGTGGATGTGCTGCAAGACGTAGAACCTACCTGCCCTGCCCTGCCCTCCC
TTCCTTATTATTCTGCTGCCAGAACATAGGTCTTGAATAAAATGGCTGGTTCTTTGTT
TCCAA
AAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 44

MALSSQIWAACLLLLL~~ASLTSGSVFPQQTGQLAELQPQDRAGARASWMPMFQRRRRDTHFPI~~
CIFCCGCCHRSKCGMCKT

Important features:

Signal peptide:

amino acids 1-24

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 58-59

N-myristoylation site.

amino acids 44-50

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 1-12

FIGURE 45

GTGGCTTCATTCAGTGGCTGACTTCCAGAGAGCAATATGGCTGGTTCCCCAACATGCCTCACCC
TCATCTATATCCTTGGCAGCTCACAGGGTCAGCAGCCTCTGGACCCGTGAAAGAGCTGGTCGGT
TCCGTTGGTGGGCCGTGACTTCCCCCTGAAGTCAAAGTAAAGCAAGTTGACTCTATTGTCTG
GACCTTCAACACAACCCCTTTGTCACCACAGCCAGAAGGGGGCACTATCATAGTGACCCAAA
ATCGTAATAGGGAGAGAGTAGACTTCCCAGATGGAGGCTACTCCCTGAAGCTCAGCAAATGAAAG
AAGAATGACTCAGGGATCTACTATGTGGGGATATACAGCTCATCCTCAGCAGCCCTCCACCCA
GGAGTACGTGCTGCATGTCTACGAGCACCTGTCAAAGCCTAAAGTCACCATGGTCTGCAGAGCA
ATAAGAATGGCACCTGTGTGACCAATCTGACATGCTGCATGGAACATGGGAAGAGGATGTGATT
TATACCTGGAAGGCCCTGGGCAAGCAGCCAATGAGTCCCATAATGGTCCATCCTCCCCATCTC
CTGGAGATGGGGAGAAAGTGTATGACCTTCATCTGCCTGCCAGGAACCCGTGAGCAGAAACT
TCTCAAGCCCCATCCTGCCAGGAAGCAGCTCTGTGAAGGTGCTGCTGATGACCCAGATTCTCCATG
GTCCTCCTGTGTCTCTGTTGGTGCCCTCCTGCTCAGTCTCTTGACTGGGCTATTCTTG
GTTTCTGAAGAGAGAGAGACAAGAAGAGTACATTGAAGAGAAGAAGAGAGTGGACATTGTCGG
AAACTCCTAACATATGCCCTATTCTGGAGAGAACACAGAGTACGACACAATCCCTCACACTAAT
AGAACAACTCTAACAGGAAGATCCAGCAAATCGTTACTCCACTGTGGAAATACCGAAAAAGAT
GGAAAATCCCCACTCACTGCTCACGATGCCAGACACACCAAGGCTATTGCCTATGAGAATGTTA
TCTAGACAGCAGTGCACTCCCCCTAAGTCTGCTCA

FIGURE 46

MAGSPTCLTLIYILWQLTGSAASGPVKELVGSVGGAVTFPLSKVKQVDSIVWTFNTTPLVTIQP
EGGTIIVTQNRNRERVDFPDGGYSLKLSKLKKNDSGIYYVGIYSSSLQQPSTQEYVLHVYEHLSK
PKVTMGLQSNKNGTCVTNLTCMEHGEEDVIYTWKALGQAANESHNGSILPISWRWGESDMTFIC
VARNPVSRNFSSPILARKLCEGAADDPDSSMVLCLLVPLLLFLVGLFLWFLKRERQEYIE
EKKRVDICRETPNICPHSGENTEYDTIPTHTNRTILKEDPANTVYSTVEIPKKMENPHSLLTMPDT
PRLFAYENVI

Important features:

Signal peptide:

amino acids 1-22

Transmembrane domain:

amino acids 224-250

Leucine zipper pattern.

amino acids 229-251

N-glycosylation sites.

amino acids 98-102, 142-146, 148-152, 172-176, 176-180, 204-208,
291-295

FIGURE 47

GGCTCGAGCGTTCTGAGCCAGGGTGACCATGACCTGCTGCGAAGGGATGGACATCCTGCAATGG
ATTCAGCCTGCTGGTTCTACTGCTGTTAGGAGTAGTTCTCAATGCGATACCTCTAATTGTCAGCT
TAGTTGAGGAAGACCAATTCTCAAACCCCATCTCTGCTTGAGTGGTGGTCCAGGAATT
ATAGGAGCAGGTCTGATGCCATTCCAGCAACAACAATGTCCTGACAGCAAGAAAAGAGCGTG
CTGCAACAACAGAACTGGAATGTTCTTCATCATTTCAGTGTGATCACAGTCATTGGTGCTC
TGTATTGCATGCTGATATCCATCCAGGCTCTCTAAAAGGTCTCTCATGTGTAATTCTCCAAGC
AACAGTAATGCCAATTGTAATTTCATTGAAAAACATCAGTGACATTCCAGAATCCTTCAA
CTTGCAGTGGTTTCAATGACTCTTGTGCACCTCCTACTGGTTCAATAAACCCACCAAGTAACG
ACACCATGGCGAGTGGCTGGAGAGCATCTAGTTCCACTTCGATTCTGAAGAAAACAAACATAGG
CTTATCCACTCTCAGTATTAGGTCTATTGCTTGGATTCTGGAGGTCTGTTGGCT
CAGTCAGATAGTCATCGGTTCTGGCTGTGTGGAGTCTCAAGCGAAGAAGTCAAATTG
TGTAGTTAATGGGAATAAAATGTAAGTATCAGTAGTTGAAAAAA

FIGURE 48

MTCCEGWTSCNGFSLLVLLLLGVVLNAIPLIVSLVEEDQFSQNPISCFEWWFPGIIGAGLMAIPA
TTMSLTARKRACCNRTGMFLSSFFSVITVIGALYCMLISIQALLKGPLMCNSPSNSNANCEFSL
KNISDIHPESFNLQWFFNDSCAPPTGFNKPTSNDTMASGWRASSFFDSEENKHRLIHFSVFLGL
LLVGILEVLFGLSQIVIGFLGCLCGVSKRRSQIV

Important features:

Transmembrane domains:

amino acids 10-31 (type II), 50-72, 87-110, 191-213

N-glycosylation sites.

amino acids 80-84, 132-136, 148-152, 163-167

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 223-227

N-myristoylation sites.

amino acids 22-28, 54-60, 83-89, 97-103, 216-222

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 207-218

TNFR/NGFR family cysteine-rich region protein.

amino acids 4-12

FIGURE 49

ATCCGTTCTCTGCGCTGCCAGCTCAGGTGAGCCCTGCCAAGGTGACCTCGCAGGACACTGGTGA
AGGAGCAGTGAGGAACCTGCAGAGTCACACAGTTGCTGACCAATTGAGCTGTGAGCCTGGAGCAG
ATCCGTGGGCTGCAGACCCCCGCCCCAGTGCCTCTCCCCCTGCAGCCCTGCCCTCGAACTGTGA
CATGGAGAGAGTGACCCCTGGCCCTTCTCCTACTGGCAGGCCTGACTGCCTTGAAGCCAATGACC
CATTGCCAATAAAGACGATCCCTTCTACTATGACTGGAAAAACCTGCAGCTGAGCGGACTGATC
TGC GGAGGGCTCCTGCCATTGCTGGATCGCGCAGTCTGAGTGGCAAATGCAAATACAAGAG
CAGCCAGAACGACAGCACAGTCCGTACCTGAGAAGGCCATCCACTCATCACTCCAGGCTTGCCA
CTACTTGCT**GAGCACAGGACTGGCCTCCAGGGATGGCCTGAAGCCTAACACTGGCCCCAGCACC**
TCCTCCCCCTGGGAGGCCTTATCCTCAAGGAAGGACTTCTCTCCAAGGGCAGGCTGTTAGGCCCC
TTCTGATCAGGAGGTTCTTATGAATTAAACTCGCCCCACCACCCCTCA

FIGURE 50

MERVTLALLLAGLTALEANDPFANKDDPFYYDWKNLQLSGLICGGLAIAGIAAVLSGKCKYKS
SQKQHSPVPEKAIPPLITPGSATTC

Important features:

Signal peptide:

amino acids 1-16

Transmembrane domain:

amino acids 36-59

N-myristoylation sites.

amino acids 41-47, 45-51, 84-90

Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7.

amino acids 54-67

FIGURE 51

GTGGACTCTGAGAAGCCCAGGCAGTTGAGGACAGGGAGAGAGAAGGCTGCAGACCCAGAGGGAGGG
AGGACAGGGAGTCGGAAGGAGGGACAGAGGAGGGCACAGAGACGCAGAGCAAGGGCGGCAAGG
AGGAGACCTGGTGGGAGGAAGACACTCTGGAGAGAGAGGGGCTGGCAGAGATGAAGTTCCAG
GGGCCCTGGCCTGCCTCCTGCTGGCCCTCTGCCTGGCAGTGGGAGGCTGGCCCTGCAGAG
CGGAGAGGAAGCACTGGACAAATATTGGGAGGGCCTGGACATGGCCTGGAGACGCCCTGA
GCGAAGGGTGGAAAGGCATTGGCAAAGAGGCCGGAGGGCAGCTGGCTCTAAAGTCAGTGAG
GCCCTGGCCAAGGGACCAGAGAAGCAGTTGGCACTGGAGTCAGGCAGGTTCCAGGCTTGGC
AGCAGATGCTTGGCAACAGGGTCGGGAAGCAGCCCAGTGTGGAAACACTGGCACGAGA
TTGGCAGACAGGCAGAAGATGTCATTGACACGGAGCAGATGCTGTCCGGCTCTGGCAGGG
GTGCCTGGCACAGTGGCTTGGAAACTTCTGGAGGCCATGGCATCTTGGCTCTCAAGGTGG
CCTTGGAGGCCAGGGCAGGGCAATCCTGGAGGTCTGGGACTCCGTGGTCCACGGATAACCCG
GAAACTCAGCAGGCAGCTTGGAAATGAATCCTCAGGGAGCTCCCTGGGTCAAGGAGGCAATGGA
GGGCCACCAAACCTTGGGACCAACACTCAGGGAGCTGTGGCCCAGCCTGGCTATGGTCAGTGAG
AGCCAGCAACCAGAATGAAGGGTCACGAATCCCCCACCATCTGGCTCAGGTGGAGGCTCCAGCA
ACTCTGGGGAGGCAGCGCTCACAGTCGGCAGCAGTGGCAGTGGCAGCAATGGTACAACAAAC
AATGGCAGCAGCAGTGGTGGCAGCAGCAGTGGCAGCAGCAGTGGCAGCAGCAGTGGGGCAGCAG
TGGCGGCAGCAGTGGTGGCAGCAGTGGCAACAGTGGTGGCAGCAGAGGTGACAGCGGCAGTGAGT
CCTCCTGGGATCCAGCACCGGCTCCTCCGGCAACCACGGTGGAGCGGGAGGAAATGG
CATAAACCCGGGTGTGAAAAGCCAGGGAAATGAAGCCGGAGCGGGGAATCTGGATTAGGG
CTTCAGAGGACAGGGAGTTCCAGCAACATGAGGGAAATAAGCAAAGAGGGAAATCGCCTCCTG
GAGGCTCTGGAGACAATTATCGGGGCAAGGGTCGAGCTGGGAGTGGAGGAGGTGACGCTGTT
GGTGGAGTCATACTGTGAACCTGAGACGTCTCCTGGATGTTAACCTTGACACTTCTGGAA
GAATTTAAATCCAAGCTGGTTCATCAACTGGGATGCCATAAACAGGACCAGAGAAGCTCTC
GCATCCCGTGACCTCCAGACAAGGAGCCACCAGATTGGATGGAGCCCCACACTCCCTCTTAA
AACACCACCCCTCATCACTAATCTAGCCCTGCCCTGAAATAACCTTAGCTGCCCAACAAA
AA
AAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 52

MKFQGPLACLLLALCLGSGEAGPLQSGEESTGTNIGEALGHGLDALSEGVGKAIGKEAGGAAGSKVS
EALGQGTREAVGTGVRQVPGFGAADALGNRVGEAAHALGNTGHEIGRQAEDVIRHGADAVRGSWQGVP
GHSGAWETSGGHGIFGSQGGLGGQGQGNGPGLGTPWVHGYPGNSAGSGFMNPQGAPWGQGGNGGPPNF
GTNTQGAVAQPGYGSVRASNQNEGCTNPPSGSGGSSNSGGSGSQSGSSGSGSNGDNNNGSSGGS
SSGSSSGSSSGSSGGSSGGSSGNSGGSRGDSGSESSWGSSTGSSGNHGGSGGGNGHKGCEKPGNE
ARGSGESGIQGFRGQGVSSNMREISKEGNRLLGGSGDNYRGQGSSWGSGGDAVGGVNTVSETSPGM
FNFDTFWKNFKSKLGFINWDAINKDQRSSRIP

Signal peptide:

amino acids 1-21

N-glycosylation site.

amino acids 265-269

Glycosaminoglycan attachment site.

amino acids 235-239, 237-241, 244-248, 255-259, 324-328, 388-392

Casein kinase II phosphorylation site.

amino acids 26-30, 109-113, 259-263, 300-304, 304-308

N-myristoylation site.

amino acids 17-23, 32-38, 42-48, 50-56, 60-66, 61-67, 64-70, 74-80,
90-96, 96-102, 130-136, 140-146, 149-155, 152-158, 155-161,
159-165, 163-169, 178-184, 190-196, 194-200, 199-205, 218-224,
236-242, 238-244, 239-245, 240-246, 245-251, 246-252, 249-252,
253-259, 256-262, 266-272, 270-276, 271-277, 275-281, 279-285,
283-289, 284-290, 287-293, 288-294, 291-297, 292-298, 295-301,
298-304, 305-311, 311-317, 315-321, 319-325, 322-328, 323-329,
325-331, 343-349, 354-360, 356-362, 374-380, 381-387, 383-389,
387-393, 389-395, 395-401

Cell attachment sequence.

amino acids 301-304

FIGURE 53

GGAGAAGAGGTTGTGGGACAAGCTGCTCCGACAGAAGGGATGTCGCTGCTGAGCCTGCCCTGG
CTGGGCCTCAGACCGGTGGCAATGTCCCCATGGCTACTCCTGCTGCTGGTTGTGGCCTGGCT
ACTCGCCCGCATCCTGGCTTGACCTATGCCTCTATAACAACTGCCGCCGGCTCCAGTGTTC
CACAGCCCCAAAACGGAACTGGTTGGGTACCTGGCCTGATCACTCCTACAGAGGAGGGC
TTGAAGGACTCGACCCAGATGTCGCCACCTATTCCCAGGGCTTACGGTATGGCTGGTCCCAT
CATCCCCTTCATCGTTTATGCCACCCCTGACACCACCGGTCTATCACCAATGCCTCAGCTGCCA
TTGCACCCAAGGATAATCTCTCATCAGGTTCTGAAGCCCTGGCTGGAGAAGGGATACTGCTG
AGTGGCGGTGACAAGTGGAGGCCACCGTCGGATGCTGACGCCGCCCTCCATTCAACATCCT
GAAGTCCTATATAACGATCTCAACAAGAGTGCACATCATGCTTGACAAGTGGCAGCACCTGG
CCTCAGAGGGCAGCAGTCGTCTGGACATGTTGAGCACATCAGCCTCATGACCTGGACAGTCTA
CAGAAATGCATCTCAGCTTGACAGCCATTGTCAAGGAGAGGCCAGTGAATATATTGCCACCAT
CTTGGAGCTCAGTGCCCTGTAGAGAAAAGAAGCCAGCATATCCTCCAGCACATGGACTTCTGT
ATTACCTCTCCATGACGGGCGGCGCTCCACAGGGCTGCCGCTGGTCATGACTCACAGAC
GCTGTCATCCGGGAGCGCGTCGCACCCCTCCCCACTCAGGGTATTGATGATTTCAAAGACAA
AGCCAAGTCCAAGACTTGGATTCATTGATGTCGCTTGCTGAGCAAGGATGAAGATGGGAAGG
CATTGTCAGATGAGGATATAAGAGCAGAGGCTGACACCTCATGTTGGAGGCCATGACACCACG
GCCAGTGGCCTCTGGTCTGTACAACCTGCGAGGCACCCAGAATACCAGGAGCGCTGCCG
ACAGGAGGTGCAAGAGCTTGTGAAGGACCGCATCTAAAGAGATTGAATGGACGACCTGGCC
AGCTGCCCTCCTGACCATGTCGTGAAGGAGAGCCTGAGGTTACATCCCCAGCTCCCTCATC
TCCCGATGCTGCACCCAGGACATTGTTCTCCAGATGGCGAGTCATCCCCAAAGGCATTACCTG
CCTCATCGATATTAGGGTCCATCACAAACCAACTGTGTGGCCGGATCCTGAGGTCTACGACC
CCTTCGCTTGACCCAGAGAAGAACAGCAAGGGAGGTACCTCTGGCTTTATTCCCTTCTCGCA
GGGCCAGGAACTGCATGGGCAGGCAGTCGCCATGGCGGAGATGAAAGTGGCCTGGCGTTGAT
GCTGCTGCACTCCGGTCTGCCAGACACACTGAGCCCCGAGGAAGCTGGAATTGATCATGC
GCGCCGAGGGCGGGCTTGGCTGCGGGTGGAGCCCTGAATGTAGGCTGCAGATTTCTGAC
CCATCCACCTGTTTGCAGATTGTCATGAATAACGGTGCTGTCAAA

FIGURE 54

MSLLSLPWGLRPVAMSPWLLLLLVGSWLLARILAWTYAFYNNCRLQCFQPPKRNFWGHLG
LITPTEEGLKDESTQMSATYSQGFTVWLGPPIIPFIVLCHPDTIRSITNASAAIAPKDNLFIRFLKP
WLGEGILLSGGDKWSRHRMLTPAFHFNILKSYITIFNKSANIMLDKWQHLASEGSSRLDMFEHI
SLMTLDSLQKCIIFSFDSHCQERPSEYIATILELSALVEKRSQHILQHMDFLYYLSHDGRRFHRAC
RLVHDFTDAVIRERRTLPTQGIDDFFKDKAKSKTLDFIGVLLSKDEDGKALSDEDIRAEADTF
MFGGHDTTASGLSWVLYNLARHPEYQERCRCQEVQELLKDRDPKEIEWDDLAQLPFLTMCVKESLR
LHPPAPFISRCCTQDIVLPDGRVIPKGITCLIDIIGVHNPTVWPDPVEVYDPFRDPENSKGRSP
LAFIPFSAGPRNCIGQAFAMAEMKVVLALMLLHFRFLPDHTEPRRKLELIMRAEGGLWLRVEPLN
VGLQ

Important features:

Transmembrane domains:

amino acids 13-32 (type II), 77-102

Cytochrome P450 cysteine heme-iron ligand signature.

amino acids 461-471

N-glycosylation sites.

amino acids 112-116, 168-172

FIGURE 55

ATCGCATCAATTGGGAGTACCATCTCCTCATGGGACCAGTGAAACAGCTGAAGCGAATGTTGA
GCCTACTCGTTGATTGCAACTATCATGGTGTGTTGCACTTACCTGTGTTCTGCCT
TTTGGTGGCATAACAAGGGACTTGCACCTATCTCTGCATTTGCAGTCTTGGCATTGACGTGG
TACAGCCTTCCTTCATACCATTGCAAGGGATGCTGTGAAGAAGTGTGTTGCCGTGTCTTGC
ATAATTCATGCCAGTTTATGAAGCTTGGAAAGGCACATGGACAGAAGCTGGTGGACAGTTT
GTAACATCTCGAAACCTCTGTCTTACAGACATGTGCCCTTATCTGCAGCAATGTGTTGCTT
GTGATTGAACATTGAGGGTTACTTTGGAAGCAACAATACATTCTCGAACCTGAATGTCAGTA
GCACAGGATGAGAAGTGGTTCTGTATCTGTGGAGTGGAAATCTCCTCATGTACCTGTTCTC
TCTGGATGTTGTCCCAGTGAATTCCCAGTAAACAAACCTATTCAACAGCAACAGCAAAAAAAA
AAA

FIGURE 56

MGPVKQLKRMFEPTRLIATIMVLLCFALTLCASFWWHNKGLALIFCILQSLALTWYSLSFIPFAR
DAVKKCFAVCLA

Important features:

Signal peptide:

amino acids 1-33

Type II fibronectin collagen-binding domain protein.

amino acids 30-72

FIGURE 57

CGGCTCGAGCTCGAGCCGAATCGGCTCGAGGGGCAGTGGAGCACCCAGCAGGCCAACATGCTCTGTCTGTGCTG
TACGTGCCGGTCATCGGGGAAGCCCAGACCGAGTTCCAGTACTTGAAGTCGAAGGGCTCCCTGCCAGCTGAAGTCC
ATTTCAAGCTCAGTGTCTCATCCCCCTCCAGGAATTCTCACCTACCGCCAGTGGAAAGCAGAAAATTGTACAAGCT
GGAGATAAGGACCTTGATGGGCAGCTAGACTTGAAGAATTGTCCATTATCTCCAAGATCATGAGAAGAAGCTGAGG
CTGGTGTAAAGATTGGACAAAAAGAATGATGGACGCATTGACGCCAGGAGATCATGCAGTCCCTGCCGGACTTG
GGAGTCAAGATATCTGAACAGCAGGCCAGAAAAAATTCTCAAGAGCATGGATAAAACGGCACGATGACCATCGACTGG
AACGAGTGGAGAGACTACCACCTCCTCCACCCCGTGGAAAACATCCCCGAGATCATCCTACTTGAAGCATTCCACG
ATCTTGATGTGGGTGAGAATCTAACGGTCCCGATGAGTCACAGTGGAGGAGGGCAGACGGGGATGTGGTGGAGA
CACCTGGTGGCAGGAGGTGGGGCAGGGGCCGTATCCAGAACCTGCACGGCCCCCTGGACAGGCTCAAGGTGCTCATG
CAGTCCATGCCCTCCCGCAGCAACACATGGCATCGTTGGCTTCACTCAGATGATTGAGAAGGAGGGGCCAGG
TCACTCTGGCAGGGCAATGGCATCACGTCTCAAATTGCCCGAATCAGCCATCAAATTATGGCTATGAGCAG
ATCAAGCGCCTTGTGGTAGTGACCAGGAGACTCTGAGGATTACAGAGAGGCTTGTGGCAGGGCTTGGCAGGGCC
ATCGCCCAGAGCAGCATCTACCAATGGAGGTCTGAAGACCCGGATGGCGCTGCCAGAACAGGCCAGTACTCAGGA
ATGCTGGACTGCCAGGAGGATCTGGCCAGAGAGGGGGTGGCGCTTCTACAAAGGCTATGTCCCCAACATGCTG
GGCATCATCCCTATGCCGGCATGCACCTTGCAGTCTACGAGACGCTCAAGAATGCCATGGCTGCAGCACTATGAGT
AACAGCGCGAACCCGGCTGTTGTGCTCCTGGCCTGTGGCACCATGTCCAGTACCTGTGCCAGCTGCCAGCTAC
CCCCCTGGCCCTAGTCAGGACCCGGATGCAGGCCAGGCCTATTGAGGGCGCTCCGGAGGTGACCATGAGCAGCCTC
TTCAAACATATCCTGCGGACCGAGGGGGCCTCGGGCTGTACAGGGGGCTGGCCCCACTCATGAAGGTATCCA
GCTGTGAGCATCAGCTACGTGGTCTACGAGAACCTGAAGATCACCTGGCGTGCAGTGCAGGTGACGGGGAGGGC
CGCCCGCAGTGGACTCGCTGATCCTGGCCGCAGCCTGGGTGTGCAGCCATCTCATTCTGTGAATGTGCCAACACT
AAGCTGTCTCGAGCCAAGCTGTAAAACCCCTAGACGCACCCGCAGGGAGGGTGGGAGAGCTGGCAGGCCAGGGCTT
GTCCTGCTGACCCCAGCAGACCCCTCTGGTCTCCAGCGAAGACCCAGGCTTCTAGGGTCCAGGGTCAGCAG
CTCCGGGCTCACATGTGTAAGGACAGGACATTTCAGCTGCAGTGCCTGCCAATAGTGAGCTGGAGCCTGGAGGCCGGCT
TAGTTCTCCATTTCACCCCTGCAGCCAGCTGTGGCCACGGCCCTGCCCTCTGGTCTGCCGTGCATCTCCCTGTGC
CCTCTGCTGCCCTGCTGTGAGGTAAGGTGGGAGGGCTACAGCCCACATCCCACCCCTCGTCCAATCCC
ATAATCCATGATGAAAGGTGAGGTACGTGGCCTCCAGGCCCTGACTCCCAACCTACAGCATTGACGCCAACCTGGC
TGTGAAGGAAGAGGAAAGGATCTGGCCTGTGGTCACTGGCATCTGAGCCCTGCTGATGGCTGGGCTCTGGGCATG
CTTGGGAGTGCAGGGGCTCGGGCTGCCCTGGCTGCACAGAAGGCAAGTGCCTGGGCTCATGGTCTGTGAGCT
GGCCTGGACCCCTGTCAAGGATGGGCCCCACCTCAGAACCAAACACTGCTCCCACTGTGGCATGAGGGCAGTGGAGCA
CCATGTTGAGGGCGAAGGGCAGAGCGTTGTGTTCTGGGAGGGAAAGGTGTTGGAGGCCCTAATTATGG
ACTGTTGGAAAAGGGTTTGTCCAGAAGGACAAGCCGGACAAATGAGCAGCTCTGTGCTTCCAGAGGAAGACGAGG
GAGCAGGAGCTGGCTGACTGCTCAGAGTCTGTGACGCCCTGGGTTCTGCAACCCAGCAGGGCGCAGC
GGGACCCAGCCCCACATTCACTTGTGCACTGCTTGGAACCTATTATTTGTATTATTAACAGAGTTATGCTCCT
AACTATTTTATAGATTGTTAATTAAATAGCTGTCAATTCAAGTCATTTCATTTTATTCATATTATGTTATGTT
GATTGTACCTCCCAAGCCGCCAGGGATGGGAGGGAGGAGGAGAAGGGGGCTTGGGGCTGCAGTCACATCT
GTCCAGAGAAATTCTTTGGACTGGAGGCAGAAAGCAGGCCAGAACGGCAGCAGCCCTGGCTCCTTCCCTGGCAG
GTTGGGAGGGCTGCCCTAGGATTCACTGGGTTGACTGGGGCTGGAGAGAGAGGGAGGAACCTCAAT
AACCTGAGGTGGAATCCAGTTATTCCCTGCCTGCAGGGTTCTTATTCACTCTTCTGAATGTCAAGGAG
TGAGGTGCCTCTACTGTGAATTGTGGTGGGGGGCTGGAGGAGGGGTGGGGCTGGCTCCGCTCCCTCCAGC
CTTCTGCTGCCCTGCTTAACAATGCCGCCAACCTGGCACCTCACGGTGCACTTCACTCCACCAAGAACATGACCTGA
TGAGGAAATCTCAATAGGATGCAAAGATCAATGCAAAATTGTTATATGAACATATAACTGGAGTCGTCAAAAG
CAAATTAAAGAAAATTGGACGTTAGAAGGTTGTCATTAAAGCAGCCTCTAATAAAAGTTGTTCAAAGCTGAAAAAA
AA

FIGURE 58

MLCLCLYVPVIGEAQTEFQYFESKGLPAELKSIFKLSVIPSQEFSTYRQWKQKIVQAGDKLDG
QLDFEEFVHYLQDHEKKLRLVFKILDKNDGRIDAQEIMQLRDLGVKISEQQAEKILKSMDKNG
TMTIDWNEWRDYHLLHPVENIPEIILYWKHSTIFDVGENLTVPDEFTVEERQTGMWWRHLVAGGG
AGAVSRTCTAPLDRLKVLMQVHASRSNNMGIVGGFTQMIREGGARSLWRGNGINVLKIAPESAIK
FMAYEQIKRLVGSQETLRIHERLVAGSLAGAIAQSSIYPMEVLKTRMALRKTGQYSGMLDCARR
ILAREGVAAFYKGYVPNMLGIIPYAGIDLAVYETLKNAWLQHYAVNSADPGVFVLLACGTMSSTC
GQLASYPLALVRTRMQAQASIEGAPEVTMSSLFKHILRTEGAFGLYRGLAPNFMKVIPAVSISYV
VYENLKITLGVQSR

Important features:

Signal peptide:

amino acids 1-16

Putative transmembrane domains:

amino acids 284-304, 339-360, 376-394

Mitochondrial energy transfer proteins signature.

amino acids 206-215, 300-309

N-glycosylation sites.

amino acids 129-133, 169-173

Elongation Factor-hand calcium-binding protein.

amino acids 54-73, 85-104, 121-140

FIGURE 59

GGAAGGCAGCGGCAGCTCCACTCAGCCAGTACCGAGATACGCTGGGAACCTTCCCCAGCCATGG
TTCCCTGGGCAGATCCTCTTCTGGAGCATAATTAGGATCATCATTATTCTGGCTGGAGCAATTG
CACTCATCATTGGCTTGGTATTCAGGGAGACACTCCATCACAGTCACTACTGTCGCCTCAGCT
GGGAACATTGGGAGGATGGAATCCTGAGCTGCACCTTGAAACCTGACATCAAACATTCTGATAT
CGTGATACAATGGCTGAAGGAAGGTGTTAGGCTGGTCATGAGTTCAAAGAAGGCAAAGATG
AGCTGTCGGAGCAGGATGAAATGTTCAGAGGCCGGACAGCAGTGTGCTGATCAAGTGATAGTT
GGCAATGCCTCTTGCCTGAAAAACGTGCAACTCACAGATGCTGGCACCTACAAATGTTATAT
CATCACTTCTAAAGGCAAGGGAAATGCTAACCTTGAGTATAAAACTGGAGCCTCAGCATGCCGG
AAGTGAATGTGGACTATAATGCCAGCTCAGAGACCTTGCGGTGTGAGGCTCCCCGATGTTCCCC
CAGCCCACAGTGGCTGGCATCCCAAGTTGACCAGGGAGCCAACCTCTCGGAAGTCTCCAATAC
CAGCTTGAGCTGAACACTTGAGAATGTGACCATGAAGGTTGTCTGTGCTCTACAATGTTACGA
TCAACAAACACATACTCCTGTATGATTGAAAATGACATTGCCAAAGCAACAGGGATATCAAAGTG
ACAGAATCGGAGATCAAAAGCGGAGTCACCTACAGCTGCTAAACTCAAAGGTTCTCTGTGTGT
CTCTTCTTCTTGCCATCAGCTGGCACTTCTGCCTCTCAGCCCTACCTGATGCTAAAATAA
GTGCCCTGGCCACAAAAAAGCATGCAAAGTCATTGTTACAACAGGGATCTACAGAACTATTCAC
CACCAAGATATGACCTAGTTATATTCTGGAGGAAATGAATTCATATCTAGAAGTCTGGAGTG
AGCAAACAAGAGCAAGAAACAAAAAGAAGCCAAGCAGAAGGCTCCAATATGAACAAGATAAAAT
CTATCTCAAAGACATATTAGAAGTTGGAAAATAATTGATGTGAACTAGACAAGTGTGTTAAGA
GTGATAAGTAAAATGCACGTGGAGACAAGTGCATCCCCAGATCTCAGGGACCTCCCCCTGCCTGT
CACCTGGGGAGTGAGAGGACAGGGATAGTGCATGTTCTTGCTCTGAATTAGTTATATGTGC
TGTAAATGTTGCTCTGAGGAAGCCCTGGAAAGTCTATCCAAACATATCCACATCTTATATTCCAC
AAATTAAGCTGTAGTATGTACCCCTAACGACGCTGCTAATTGACTGCCACTCGCAACTCAGGGCG
GCTGCATTTAGTAATGGGTCAAATGATTCACTTTTATGATGCTTCAAAGGTGCCTGGCTTC
TCTTCCCAACTGACAAATGCCAAAGTTGAGAAAATGATCATATTAGCATAAACAGAGCAGT
CGGGGACACCAGTTATAAATAACTGAGCACCTCTTTAAACAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 60

MASLGQILFWSIISIIILAGAIALIIGFGISGRHSITVTTVASAGNIGEDGILSCTFEPDIKLS
DIVIQWLKEGVGLGLVHEFKEGKDELSEQDEMFRGRTAVFADQVIVGNASLRLKNVQLTDAGTYKC
YIITSKGKGNANLEYKTGAFSMPEVNVDYNASSETLRCEAPRWFPQPTVVWASQVDQGANFSEVS
NTSFELNSENVTMKVSVLYNVTINNTYSCMIENDIAKATGDIKVTESEIKRRSHLQLLNSKASL
CVSSFFAISWALLPLSPYIMLK

Important features:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 258-281

N-glycosylation sites.

amino acids 112-116, 160-164, 190-194, 196-200, 205-209, 216-220,
220-224

N-myristoylation sites.

amino acids 52-58, 126-132, 188-194

FIGURE 61

TGACGTAGAATCACCCATGGCCAGCTATCCTACCGCAGGGCTGCCAGGAGCTGCAGGACAAG
CACCAAGGAGCCCCTCCGGTAGCTACTACCCTGGACCCCCAATAGTGGAGGGCAGTATGGTAGT
GGGCTACCCCTGGTGGTATGGGGTCTGCCCTGGAGGGCTTATGGACCACCAGCTGG
TGGAGGGCCCTATGGACACCCCAATCCTGGATGTTCCCTCTGGAACCTCCAGGAGGACCATATG
GCGGTGCAGCTCCGGGGCCCTATGGTCAGCCACCTCCAAGTCCCTACGGTGCCAGCAGCCT
GGGCTTATGGACAGGGTGGGCCCTCCAATGTGGATCCTGAGGCCTACTCCTGGTCCAGTC
GGTGGACTCAGATCACAGTGGCTATATCTCCATGAAGGAGCTAAAGCAGGCCCTGGTCAACTGCA
ATTGGTCTTCATTCAATGATGAGACCTGCCTCATGATGATAAACATGTTGACAAGACCAAGTCA
GGCCGCATCGATGTCTACGGCTTCTCAGCCCTGGAAATTCCAGCAGTGGAAAGAACCTCTT
CCAGCAGTATGACCGGGACCGCTCGGCTCCATTAGCTACACAGAGCTGCAGCAAGCTGTCCC
AAATGGGCTACAACCTGAGCCCCAGTTCACCCAGCTTCTGGTCTCCGCTACTGCCACGCTCT
GCCAATCCTGCCATGCAGCTGACCGCTTCATCCAGGTGTGCACCCAGCTGCAGGTGCTGACAGA
GGCCTCCGGAGAAGGACACAGCTGTACAAGGCAACATCCGGCTCAGCTCGAGGACTTCGTCA
CCATGACAGCTCTCGGATGCTATGACCCAAACCATCTGTGGAGAGTGGAGTGCACCAAGGGACCTT
TCCTGGCTTCTTAGAGTGAGAGAAGTATGTGGACATCTCTCTTCCCTGTCCCTCTAGAAGAAC
ATTCTCCCTGCTTGATGCAACACTGTTCAAAGAGGGTGGAGAGTCCTGCATCATGCCACCA
AATAGTGAGGACCGGGCTGAGGCCACACAGATAGGGCCTGATGGAGGAGAGGATAGAAGTTGA
ATGTCCTGATGCCATGAGCAGTTGAGTGGCACAGCCTGGCACCAGGAGCAGGTCTGTAATGG
AGTTAGTGTCCAGTCAGCTGAGCTCCACCCCTGATGCCAGTGGTAGTGTCTGACCGCTGTTACC
GTTAGTACCTGTGTTCCCTCACCAGGCCATCCTGTCAAACGAGCCATTCTCCAAAGTGGAAAT
CTGACCAAGCATGAGAGAGATCTGTCTAGCTGCTGGGACCTGGCTGGATTCTGCCACACCCATAAAT
CCTTGTGTGTTAACTCTAGCTGCCTGGGCTGCCCTGCTCAGACAAATCTGCTCCCTGGCAT
CTTGGCCAGGCTCTGCCCTGCAGCTGGACCCCTCACTGCCTGCCATGCTGCTCGGCT
TCAGTCTCCAGGAGACAGTGGTCACCTCCCTGCCAATACTTTTAATTGCATTTTTC
ATTTGGGCCAAAGTCCAGTGAAATTGTAAGCTCAATAAAAGGATGAAACTCTGA

FIGURE 62

MASYPYRQGCPGAAGQAPGAPPGSYYPGPPNSGGQYGSGLPPGGYGGPAPGGPYGPPAGGGPYG
HPNPGMFPSTPGGPYGGAAPGGPYGQPPPSSYGAQQPGLYQGGAPPNVDPEAYSWFQSVDSDH
SGYISMKELKQALVNCNWSSFNDETCLMMINMFDTKSGRIDVYGFSAWKFQQWKNLFQQYDR
DRSGSISYTELQQALSQMGYNLSPQFTQLLVSRYCPRSANPAMQLDRFIQVCTQLQVLTEAFREK
DTAVQGNIRLSFEDFVTMTASRML

Important features of the protein:

Signal peptide:

amino acids 1-19

N-glycosylation site.

amino acids 147-150

Casein kinase II phosphorylation sites.

amino acids 135-138, 150-153, 202-205, 271-274

N-myristoylation sites.

amino acids 9-14, 15-20, 19-24, 33-38, 34-39, 39-44, 43-48, 61-66, 70-75, 78-83, 83-88, 87-92, 110-115

FIGURE 63

CAGG**A**TGCAGGGCCCGTGGCAGGGAGCTGCCTCCTCTGGGCCTGCTCCTGGTCTGTCTTCATC
TCCCAGGCCTCTTGCCCCGGAGCATCGGTGTTGGAGGAGAAAGTTCCAAAACCTCGGGACC
AACTGCCTCAGCTCGGACAACCTCCTCCACTGGCCCCTCTAACACTGAACATCCGAGCCCCGC
TCTGGACCCTAGGTCTAATGACTTGGCAAGGGTTCCTCTGAAGCTCAGCGTGCCTCCATCAGATG
GCTTCCCACCTGCAGGAGGTTCTGCAGTGCAGAGGTGGCCTCCATCGTGGGGCTGCCATG
GATTCTGGCCCCCTGAGGATCCTTGGCAGATGATGGCTGCTGCAGGACCGCCTGGGG
AGCGCTGCCTGAAGAACTCTTACCTCTCCAGTGCTGCCAGGCTCGCTCCGGCAGTGGCCCTT
TGCCTGGGAGTCTCTCCGATGCCACAGGCCTCTCACCTGAGGTTCACTCCTCCACCAGGAC
TCGGAGTCCAGACGACTGCCCGTTCTAACATTCACTGGGAGCCGGGGAAAAATCCTTCCAAACG
CCCTCCCTGGTCTCTCATCCACAGGGTTCTGCCTGATCACCCCTGGGTACCCCTGAATCCCAGTG
TGTCTGGGAGGTGGAGGCCCTGGACTGGTTGGGAACGAGGCCATGCCACACCCCTGAGGGA
ATCTGGGTATCAATAATCAACCCCCAGGTACCACTGGGAAATATTAATCGGTATCCAGGAGG
CAGCTGGGAAATATTAATCGTATCCAGGAGGCAGCTGGGAAATATTAATCGGTATCCAGGAG
GCAGCTGGGAAATATTCTACCTATACCCAGGTATCAATAACCCATTCCCTCTGGAGTTCTCCGC
CCTCCTGGCTCTTCTTGGAACATCCCAGCTGGCTCCCTAATCCTCCAAGCCCTAGGTTGCAGTG
GGG**C**AGAGCACGATAGAGGGAAACCCAACATTGGAGTAGAGTCTGCTCCGCCCTGCTG
TGTGGGCTCAATCCAGGCCCTGTTAACATGTTCCAGCACTATCCCCACTTTCAAGTGCCTCCCC
TGCTCATCTCCAATAAAATAAAAGCACTTATGAAAAA
AA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 64

MQGRVAGSCAPLGLLVCLHLPGLFARSIGVVEEKVSQNFQGTNLPOQPSSTGPSNSEHPQPAL
DPRSNDLARVPLKLSVPPSDGFPPAGGSAVQRWPPSWGLPAMDSWPPEDPWQMMAAAEDRLGEA
LPEEELSYLSSAAALAPGSGPLPGESSPDATGLSPEASLLHQDSESRRLPRSNSLGAGGKILSQRP
PWSLIHRVLPDHPWGTLNPSVSWGGGGPGTGWGTRPMPHPEGIWGINNQPPGTSWGNINRYPGGS
WGNINRYPGGSWGNINRYPGGSWGNIHLYPGINNPFPPGVLRPPGSSWNIPAGFPNPPSPRLQWG

Important features of the protein:

Signal peptide:

amino acids 1-26

Casein kinase II phosphorylation sites.

amino acids 56-59, 155-158

N-myristoylation sites.

amino acids 48-53, 220-225, 221-226, 224-229, 247-252, 258-263,
259-264, 269-274, 270-275, 280-285, 281-286, 305-310

FIGURE 65

AAGGAGAGGCCACCGGGACTTCAGTGTCTCCTCCATCCCAGGAGCGCAGTGGCCACTATGGGTC
TGGGCTGCCCTTGTCCCTCCTTGCACCTCCTGGCAGCTCACATGGAACAGGGCCGGGTATGA
CTTGCAACTGAAGCTGAAGGAGTCTTCTGACAAATTCCCTCATGAGTCCAGCTCCTGGAA
TTGCTTGAAAAGCTCTGCCTCCTCCATCTCCCTCAGGGACCAGCGTACCCCTCACCATGC
AAGATCTCAACACCATGTTGTCGCAACACATTGACAGCCATTGAAGCCTGTCTTGGCCC
GGGCTTTGGGCCGGGATGCAGGAGGCAGGCCCGACCCTGTCTTCAGCAGGCCACCCTC
CTGAGTGGCAATAAATAAATTGGTATGCTG

FIGURE 66

MGSGPLVLLTLLGSSHGTGPGMTLQLKLKESFLTNSSYESSFELLEKLCLLHLPSGTSVTL
HHARSQHHVVCNT

Important features:

Signal peptide:

amino acids 1-19

N-glycosylation site.

amino acids 37-41

N-myristoylation sites.

amino acids 15-21, 19-25, 60-66

FIGURE 67

ACGGACCGAGGGTTCGAGGGAGGGACACGGACCAGGAACCTGAGCTAGGTCAAAGACGCCGGC
CAGGTGCCCGTCGCAAGTGCCCTGGCCGGAGATGCGGTAGGAGGGCGAGCGCGAGAACCCCC
TTCCTCGCGCTGCCAACCCGCCACCCAGCCATGGCGAACCCGGCTGGGCTGCTTCTGGCG
CTGGGCCTGCCGTTCTGCTGGCCGCTGGGCGAGCCTGGGGCAAATACAGACCACTCTGC
AAATGAGAATAGCACTGTTGCCTTCATCCACCAGCTCCAGCTCCGATGGCAACCTGCGTCCGG
AAGCCATCACTGCTATCATCGTGGTCTTCTCCCTTGGCTGCCTGCTCCTGGCTGTGGGCTG
GCACTGTTGGTGCAGAAGCTCGGAGAAGCGGCAGACGGAGGGCACCTACCGGCCAGTAGCGA
GGAGCAGTTCTCCATGCAGCCGAGGCCGGCCCTCAGGACTCCAAGGAGACGGTGCAGGGCT
GCCTGCCATCAGGTCCCCTCCTGCATCTGTCTCCCTCATTGCTGTGACCTTGGGAAA
GGCAGTGCCCTCTGGCAGTCAGATCCACCCAGTGCTTAATAGCAGGGAAAGGTACTTCAA
AGACTCTGCCCTGAGGTCAAGAGAGGATGGGCTATTCACTTTATATATTTATATAAAATTAG
TAGTGAGATGTAAAAAAAAAAAAAA

FIGURE 68

MANPGLGLLLALGLPFLARWGRAWGQIQTTSANENSTVLPSSSTSSSDGNLRPEAITAIIVFS
LLAALLLAVGLALLVRKLREKRQTEGTYRPSSEEQFSHAAEARAPQDSKETVQGCLPI

Important features:

Signal peptide:

amino acids 1-19

Transmembrane domain:

amino acids 56-80

N-glycosylation site.

amino acids 36-40

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 86-90

Tyrosine kinase phosphorylation site.

amino acids 86-94

N-myristoylation sites.

amino acids 7-13, 26-32

FIGURE 69

GCCAGGAATAACTAGAGAGGAACAATGGGGTTATTCAGAGGTTTGTTCCTCTAGTTCTGTGCCCTGCTGCACCAAG
TCAAATACTTCCTTCATTAAGCTGAATAATAATGGCTTGAAGATATTGTCAATTGTTAGATCCTAGTGTGCCAGAA
GATGAAAAAAATAATTGAACAAATAGAGGATATGGTGAACCTACAGCTTACGTACCTGTTGAAGGCCACAGAAAAAGA
TTTTTTCAAAAATGTATCTATTAATTCTGAGAATTGGAAGGAAATCCTCAGTACAAAAGGCCAAACATGAA
AACCATAAACATGCTGATGTTAGTGCACCACTACACTCCCAGGTAGAGATGAACCATACACCAAGCAGTCACA
GAATGTGGAGAGAAAGGCGAATACATTCACTCACCCCTGACCTTACTTGGAAAAAAACAAATGAATATGGACCA
CCAGGCAAACGTGTTGTCCATGAGTGGGCTCACCTCCGGTGGGAGTGTGTTGATGAGTACAATGAAGATCAGCCTTC
TACCGTCTAAGTCAAAAAAATCGAAGCAACAAGGTGTTCCGCAGGTATCTCTGGTAGAAATAGAGTTATAAGTGT
CAAGGAGGCAGCTGTCTTAGTAGGCATGCAGAATTGATTCTACAACAAAATGTATGGAAAAGATTGTCAATTCTT
CCTGATAAAAGTACAAACAGAAAAGCATCCATAATGTTATGCAAAGTATTGATTCTGTTGAATTGTAAAGAA
AAAACCCATAATCAAGAAGCTCCAAGCCTACAAAACATAAAAGTCAAAAGTCAATTAGTAGTACATGGGAGGTGATTAGCAAT
TCTGAGGATTTAAAAACACCATACCCATGGTGAACACCACCTCCACCTGCTTCTCATGCTGAAGATCAGTC
AGAATTGTGTGCTTAGTCTTGATAAGTCTGGAAAGCATGGGGTAAGGACCGCCTAAATCGAATGAATCAAGCAGCA
AAACATTTCCTGCTGCAGACTGTTGAAAATGGATCCTGGGTGGGATGGTCACTTGATAGTACTGCCACTATTGTA
AATAAGCTAATCCAATAAAAGCAGTGAAGAAACACACTCATGGCAGGATTACCTACATATCCTCTGGGAGGA
ACTTCCATCTGCTCTGGAATTAAATATGCATTCAGGTGATTGGAGAGCTACATCCCAACTCGATGGATCCGAAGTA
CTGCTGCTGACTGATGGGAGGATAACACTGCAAGTCTTGTATTGATGAAGTGAACAAAGTGGGCCATTGTTCAT
TTTATTGCTTGGGAGAGCCTGCTGATGAAGCAGTAATAGAGATGAGCAAGATAACAGGAGGAAGTCATTTATGTT
TCAGATGAAGCTCAGAACATGGCCTCATTGATGCTTGGGCTCTACATCAGGAAATACTGATCTCCAGAAG
TCCCTCAGCTCGAAAGTAAGGGATTAACACTGAATAGTAATGCCCTGGATGAACGACACTGTCATAATTGATAGTACA
GTGGAAAGGACACGTTCTTCATCACATGGAACAGTCTGCCTCCAGTATTCTCTGGATCCAGTGGAAACA
ATAATGGAAAATTTCACAGTGGATGCAACTTCCAAAATGGCCTATCTCAGTATTCCAGGAACCTGCAAAGGTGGCACT
TGGGCATACAATCTTCAAGCCAAAGCGAACCCAGAAACATTAACACTATTACAGTAACCTCTCAGCAGCAAATTCTCT
GTGCCCTCAATCACAGTGAATGCTAAAGAATAAGGACGTAAACAGTCTCCCAGCCAATGATTGTTACGCAGAA
ATTCTACAAGGATATGTAACCTGTTGGAGGCAATGTGACTGCTTCATTGAATCACAGAATGGACATACAGAAGTT
TTGGAACCTTGGATAATGGTGCAGGCGCTGATTCTTCAAGAATGATGGAGTCACCCAGGTATTTACAGCATAT
ACAGAAAATGGCAGATATAGCTAAAAGTCTGGGCTCATGGAGGAGCAAACACTGCCAGGCTAAAATTACGGCTCCA
CTGAATAGAGCCCGTACATACCAGGCTGGGTAGTGAACGGGAAATTGAAGCAAACCGCCAAGACCTGAAATTGAT
GAGGATACTCAGACCACCTGGAGGATTTCAGCCGAACAGCATCCGGAGGTGCATTGTTGATCACAGTCCAGC
CTTCCCTGCTGACCAATACCCACCAAGTCACAGACCTTGATGCCACAGTCATGAGGATAAGGATTATTCT
ACATGGACAGCACCAGGAGATAATTGATGTTGGAAAAGTCAACGTTATATCATAAGAATAAGTGCAGTATTCT
GATCTAAGAGACAGTTGATGATGCTCTCAAGTAAATACTACTGATCTGTCACCAAGGAGGCAACTCCAAGGAA
AGCTTGCATTAAACAGAAAATATCTCAGAAGAAAATGCAACCCACATATTGATGCTTAAAGTATAGATAAA
AGCAATTGACATCAAAGTATCCACATTGCAACAGTAACCTTGTTATCCCTCAAGCAAATCCTGATGACATTGAT
CCTACACCTACTCCTACTCCTACTCCTGATAAAAGTCATAATTCTGGAGTTAATATTCTACGCTGGTATTG
TCTGTGATTGGCTGTTGTAATTGTTACCTTATTAAAGTACCCATTTGAACCTTAACGAAGAAAAATCTTC
AAGTAGACCTAGAAGAGGTTAAAAACAAATGTAAGTAAAGGATATTCTGAATCTTAAATTCCAT
GTGTGATCATAAAACTCATAAAAATAATTAAAGATGTCGGAAAAGGATACCTTGATTAAATAAAAACACTCATGGATA
TGTAAAAACTGTCAAGATTAAATGTTCAATTGTTATTGTTATTGTAAGAAATAGTGTGAACAAAG
ATCCTTTTCAACTGATACCTGGTTGTATATTGATGCAACAGTTCTGAAATGATATTCAAATTGCA
GAAATTAAAATCATCTATGAGTAGTCAAATACAAGTAAAGGAGGAGCAAATAACACATTTGGAAAAAA
AA

FIGURE 70

MGLFRGFVFLVLCLLHQSNSTSIKLNNGFEDIVIVIDPSVPEDEKIIEQIEDMVTTASTYLFE
ATEKRFKKNVSILIPENWKENPQYKRPKHENKHADVIVAPTLPGRDEPYTKQFTECGEKG
IHFTPDLGGKQNEYGPPGKLFVHEWAHLRGVFDEYNEDQPFYRAKSKKIEATRCSAGISGRN
RVYKCQGGSCLSACRIDSTTKLYGKDCQFFPDKVQTEKASIMFMQSIDSVVEFCNEKTHNQEAP
SLQNIKCNFRSTWEVISNSEDFKNTIPMVTPPPPVFSLLKISQRIVCLVLDKSGSMGGKDRLNR
MNQAAKHFLQLQTVENGSWGMVHFSTATIVNKLIQIKSSDERNTLMAGLPTYPLGGTSICSGIK
YAFQVIGELHSQDGSEVLLTDGEDNTASSCIDEVKQSGAIVHFIALGRAADEAVIEMSKITGG
SHFYVSDEAQNNGLIDAFGALTSGNTDLSQKSLQLESKGLTLNSNAWMNDVIIDSTVGKDTFFL
ITWNSLPPSISLWDPSGTIMENFTVADATSKMAYLSIPGTAKVGTWAYNLOAKANPETLTITVTSR
AANSSVPPITVNAKMNDVNSFPSPMIVYAEILQGYVPVLGANVTAFIESQNGHTEVLELLDNGA
GADSFKNDGVSYRYFTAYTENGRYSLKVRAGGANTARLKLRRPLNRAAYIPGWVVNGEIEANPP
RPEIDEDTQTTLEDFSRTASGGAFVVSQVPSLPLPDQYPPSQITDLDATVHEDKIIILTWTAPGDN
FDVGKVQRYIIRISASILDLRDSFDDALQVNTTDLSPKEANSKESFAFKPENISEENATHIFIAI
KSIDKSNLTSKVSNIAQVTLFIPQANPDDIDPTPTPTPTDKSHNSGVNISTLVLSVIGSVVI
VNFILESTTI

Signal peptide:

amino acids 1-21

Putative transmembrane domains:

amino acids 284-300, 617-633

Leucine zipper pattern.

amino acids 469-491, 476-498

N-glycosylation site.

amino acids 20-24, 75-79, 340-344, 504-508, 542-546, 588-592, 628-632, 811-815, 832-836, 837-841, 852-856, 896-900

FIGURE 71

CTCCTTAGGTGAAACCCCTGGGAGTAGAGTACTGACAGCAAAGACCGGGAAAGACCATACTGCCCCGGCAGGGGTGA
CAACAGGTGTCATCTTTGATCTCGTGTGGCTGCCTTCCTATTCAAGGAAAGACGCCAAGGTAATTTGACCCA
GAGGAGCAATGATGTAGCCACCTCCTAACCTCCCTCTTGAACCCCCAGTTATGCCAGGATTACTAGAGAGTGTC
ACTCAACCAGCAAGCGGCTCCTCGGCTTAACCTGTGGTGGAGGAGAGAACCTTGTGGGCTGCGTCTCTTAGCA
GTGCTCAGAAGTGAECTGCCTGAGGGTGGACAGAAGAAAGGAAAGGCTCCCTCTGCTGTGGCTGCACATCAGGAA
GGCTGTGATGGGAATGAAGGTGAAACTTGGAGATTCACTTCAGTCATTGCTCTGCCTGCAAGGATCATCCTTAAA
AGTAGAGAAGCTGCTCTGTTGGGTTAACCTCAAGAGGCAGAACCTGTTCTAGAAGGAAATGGATGCAAGCAGCTC
CGGGGCCCCAACGCATGCTTCCTGTGGCTAGGCCAGGGAAGGCCCTCCGTGGGGCCCGCTTGAGGGATGCC
ACCGGTTCTGGACGCATGGCTGATTCTGAATGATGGTCTGCCGGGCTGCTTGCCTGGGAGTTCCCGGGTGGT
GTTTGCTGGTCTCTGCTGTGCTATCTGCTGTACATGTTGGCCTGACCCCAAAGGGTACGAGGAGCAG
CTGGCACTGCCAGGGCAACAGCCCCACGGGAAGGAGGGTACCAAGGCCGTCCTCAGGAGTGGGAGGAGCAGCAC
CGCAACTACGTGAGCAGCCTGAAGCGGCAGATCGCACAGCTCAAGGAGGAGCTGCAGGAGAGGAGTGAGCAGCTCAGG
AATGGGAGTACCAAGCCAGCGATGCTGTGGCCTGGGTCTGGACAGGAGCCCCCAGAGAAAACCCAGGCCGACCTC
CTGGCCTTCCTGCACTCGCAGGTGGACAAGGCAGAGGTGAATGCTGGCGTCAAGCTGCCACAGAGTATGCAGCAGTG
CCTTCGATAGCTTACTCTACAGAAGGTGTACAGCTGGAGACTGCCCTACCCGCCACCCGAGGAGAACGCTGTG
AGGAAGGACAAGCGGGATGAGTTGGGAAGCATTGAATCAGCCTGGAGACCCCTGAACAACTCTGCAAGAGAACAGC
CCCAATCACCGCTTACACGGCCTGATTCAAGGGATCTACCGAACAGAAAGGGACAAAGGGACATTGTAT
GAGCTCACCTCAAAGGGGACCACAAACACGAATTCAAACGGCTCATCTTATTCGACCATTAGCCCCATCATGAAA
GTGAAAAATGAAAAGCTCAACATGGCAACACGGCTTATCAATGTTATCGTGCCTCTAGCAAAAAGGGTGGACAAGTC
CGGCAGTTATGCAAGATTCAAGGGAGATGTGATTGAGCAGGATGGGAGAGTCATCTACTGTTGTTACTTGGG
AAAGAAGAAATAATGAAGTCAAAGGAATACTTGAAAACACTTCAAAGCTGCCAACTTCAGGAACCTTACCTTCATC
CAGCTGAATGGAGAATTTCAGGGAAAGGGACTTGATGTTGGAGGCCGCTCTGGAGGGAAACGTCCTTCT
TTTTCTGTGATGTTGGACATCTACTCACATGAAATTCTCAATACGTGTAGGCTGAATACACAGCCAGGGAAAG
GTATTTATCCAGTTCTTCAGTCAGTACAATCCTGGCATAATATACGGCCACCATGATGCAGTCCCTCCCTGGAA
CAGCAGCTGGTCATAAAGAAGGAAACTGGATTGGAGAGACTTTGGATTGGATGACGTGTCAGTATCGGTCA
TTCATCAATATAGTGGGTTGATCTGGACATCAAAGGCTGGGCGGAGAGGATGTGCACCTTATCGCAAGTATCTC
CACAGCAACCTCATAGTGGTACGGACGCCTGCGAGGACTCTCCACCTCTGGCATGAGAAGCGCTGCATGGACAG
CTGACCCCCGAGCAGTACAAGATGTGCACTGCAAGGCCATGAAAGGAGCTGGCAGTGGCATGCT
GTGTTCAAGGACAGATAGAGGCTCACCTCGAACAGAAACAGAACAGAACAGTAGCAAAAAACATGAACTCCAGA
GAAGGATTGTGGAGACACTTTCTTCTTTGCAATTACTGAAAGTGGCTGCAACAGAGAAAAGACTTCCATAAA
GGACGACAAAAGAATTGGACTGATGGGTAGAGATGAGAAAGCCTCCGATTCTCTGTTGGGCTTTACAACAGA
AATCAAAATCTCGCTTGCCTGCAAAGTAACCCAGTTGCACCCCTGTGAAGTGTCTGACAAAGGCAGAATGCTTGT
AGATTATAAGCTAACTGGTGTGGAGGTTTGATGGTGTAACTACACTGAGACCTGTTGTGCTCATGAA
AATATTGATTTAAGAGCAGTTGTAAGGAAATTCTGATGAAAGGCAAGCATATTCTCTCATATGAATGA
GCCTATCAGCAGGGCTCTAGTTCTAGGAATGCTAAATATCAGAAGGCAGGAGAGGAGATAGGCTTATTATGATACT
AGTGAGTACATTAAGTAAATGGACCAGAAAAGAAAAGAACATAAATATCGTGTCTATTCTCCCAAGAT
TAACCAAAATAATCTGCTTATCTTGGTGTCTTTAACTGCTCCGTTTCTTATTTAAATGACT
TTTTCTCCCTGTGAGTTAGTCTGCTTATTAAATTACCACTTGCACGCCCTACAAGAGAGCACAAGTGGCCTAC
ATTTTTATTTTAAAGAGATACTTGAGATGCAATTGAGAACCTTCAGTCACGCAATTGAGACATAGGAAGGAATGGTTGTACT
CCAAGGACATGCCAAATGCTGATTCTGCAAGGACTGAATGTCAGGCAATTGAGACATAGGAAGGAATGGTTGTACT
AATACAGACGTACAGATACTTCTGAAAGAGTATTTGCAAGAGGAGCAACTGAACACTGGAGGAAAGAAAATGAC
ACTTCTGCTTACAGAAAAGGAAACTCATTCAAGACTGGTATCTGATGTCAGCTAAAGTCAGAAACCACATT
CTCCTCAGAAGTAGGGACCGCTTCTACCTGTTAAATAACCAAGTATACCGTGTGAACCAAACAAATCTCTTTC
AAAACAGGGTGTCTCCTGGCTCTGGCTTCCATAAGAAGAAATGGAGAAAATATATATATATATATATG
GAAAGATCAATCCATCTGCCAGAACATCTAGGGATGGAAGTTTGCTACATGTTATCCACCCAGGCCAGGTGGAG
TAAGTAACTGTTTAAAGCAGTTACTCAATACCAAGATGCTCTGAAATTGCTATTGCTTATTACCA
CAAACATTTTAAAGGAAATACAGTTAACATAGAGTGGTTCTTCATTGAGTAAAGGCTTGTACAGTAAACTCATTGTTAAAGCTTCAAGAAC
ATGCAATGAGCTAAATTATCTCTTGAGTCCTGCTGTGCTCACAGTAAACTCATTGTTAAAGCTTCAAGAAC
ATTCAAGCTGTGGTGTGTTAAAAATGCATTGATTGATTGACTGGTAGTTATGAAATTAAATTAAAACACAGG
CCATGAATGGAAGGTGGTATTGCACAGCTAATAAAATATGATTGTGGATATGAA

FIGURE 72

MMMVRRGLLAWISRVVVLLVLLCAISVLYMLACTPKGDEEQLALPRANSPTGKEGYQAVLQEWE
EQHRNYVSSLKRQIAQLKEELQERSEQLRNGQYQASDAAGLGLDRSPPEKTQADLLAFLHSQVDK
AEVNAGVKLATEYAAVPFDSFTLQKVYQLETGLTRHPEEKPVRKDKRDELVEAIESALETLNNPA
ENSPNHRPYTASDFIEGIYRTERDKGTLYELTFKGDHKHEFKRLILFRPFSPIMKVKNEKLNMAN
TLINVIVPLAKRVDKFRQFMQNREMCEQDGRVHLTUVYFGKEEINEVKGILENTSKAANFRNF
TFIQLNGEFSRGKGLDVGARFWKGSNVLLFFCDVDIYFTSEFLNTCRLNTQPGKKVYPVLFQSQY
NPGIIYGHDAVPPLEQQLVIKKETGFWRDFGFGMTQCYRSDFINIGGFDL DIKGWGGEDVHLYR
KYLHSNLIVVRTPVRGLFHLWHEKRCMDELTPEQYKMCMQSKAMNEASHGQLGMLVFRHEIEAHL
RKQKQKTSSKKT

Important features:

Signal peptide:

amino acids 1-27

N-glycosylation sites.

amino acids 315-319, 324-328

N-myristoylation sites.

amino acids 96-102, 136-142, 212-218, 311-317, 339-345, 393-399

Amidation site.

amino acids 377-381

FIGURE 73

GAGACTGCAGAGGGAGATAAAGAGAGAGGGCAAAGAGGGCAGCAAGAGATTGTCCTGGGATCCA
GAAACCATGATACCTACTGAACACCGAATCCCTGGAAGCCCACAGAGACAGAGACAGCAAGA
GAAGCAGAGATAAATACACTCACGCCAGGAGCTCGCTCGCTCTCTCTCTCACTCCTC
CCTCCCTCTCTGCCTGTCCTAGTCCTCTAGTCCTCAAATTCCCAGTCCCCTGCACCCCTTC
CTGGGACACTATGTTCTCCGCCCTCCTGCTGGAGGTGATTGGATCCTGGCTGCAGATGGG
GTCAACACTGGACGTATGAGGGCCCACATGGTCAGGACCATTGCCAGCCTTACCCGTAGTGT
GGAAACAATGCCAGTCGCCATCGATATTCAAGACAGACAGTGTGACATTGACCCGTAGTGT
TGCTCTGCAGCCCCACGGATATGACCAGCCTGGCACCGAGCCTTGGACCTGCACAACAATGCC
ACACAGTGCAACTCTCTGCCCTCACCCGTATCTGGGTGGACTTCCCCGAAAATATGTAGCT
GCCAGCTCCACCTGCACTGGGTCAAGAAAGGATCCCCAGGGGGTCAGAACACCAAGATCAACAG
TGAAGCCACATTGCAAGAGCTCCACATTGTACATTATGACTCTGATTCTATGACAGCTTGAGTG
AGGCTGCTGAGAGGCCTCAGGGCCTGGCTGTCCCTGGCATCCTAATTGAGGTGGTGAGACTAAG
AATATAGCTTATGAACACATTCTGAGTCATTGCATGAAGTCAGGCATAAAGATCAGAAGACCTC
AGTGCCTCCCTCAACCTAACAGAGAGCTGCTCCCAAACAGCTGGGCAGTACTCCGCTACAATG
GCTCGCTACAACCTCCCTGCTACCAGAGTGTGCTCTGGACAGTTTTATAGAAGGTCCAG
ATTCAATGGAACAGCTGGAAAAGCTTCAGGGACATTGTTCTCACAGAACAGGAGCCCTCTAA
GCTTCTGGTACAGAACTACCGAGCCCTCAGCCTCTCAATCAGCGCATGGCTTGCTTCA
TCCAAGCAGGATCCTCGTATACCACAGGTGAAATGCTGAGTCTAGGTGAGGAATCTGGTTGGC
TGTCTGCTGCTTCTCCTGGCTGTTATTCATTGCTAGAAAGATTGGAAGAACAGGCTGGAAA
CCGAAAGAGTGTGGCTTCACCTCAGCACAAAGCCACGACTGAGGCATAAATCCTCTCAGATAC
CATGGATGTGGATGACTCCCTCATGCCCTACAGGAAGCCTCTAAATGGGTGAGGATCTGG
CCAGAAACACTGTAGGAGTAGTAAGCAGATGTCCCTCCCTGGACATCTCTAGAGAGGAAT
GGACCCAGGCTGTCATTCCAGGAAGAACTGCAGAGCCTCAGCCTCTCAAACATGTAGGAGGAA
ATGAGGAAATCGCTGTTGTTAATGCAGAGANCAAACCTCTGTTAGTTGCAGGGAAAGTTGGG
ATATACCCCAAAGTCCCTACCCCTCACTTTATGCCCTTCCCTAGATATACTGCGGGATCT
CTCCTTAGGATAAAGAGTTGCTGTTGAAGTTGATATTTGATCAATATATTGAAATTAAAG
TTCTGACTTT

FIGURE 74

MLFSALLLEVIWILAADGGQHWTYEGPHGQDHWPASYPECNNAQSPIDIQTDSVTFPDLPALQ
PHGYDQPGTEPLDLHNNGHTVQLSLPSTLYLGGLPRKYVAAQLHLHWGQKGSPGGSEHQINSEAT
FAELHIVHYDSDSYDSLSEAAERPQGLAVLGILIEVGETKNIAYEHILSHLHEVRHKDQKTSVPP
FNLRELLPKQLGQYFRYNGSLTPPCYQSVLWTVFYRRSQISMEQLEKLQGTLFSTEEPSKLLV
QNYRALQPLNQRMVFASFIQAGSSYTTGEMLSLGVGILVGCLLLLAVYFIARKIRKKRLENRKS
VVFTSAQATTEA

Important features of the protein:

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 291-310

N-glycosylation site.

amino acids 213-216

Eukaryotic-type carbonic anhydrases proteins

amino acids 197-245, 104-140, 22-69

FIGURE 75

TGCCGCTGCCGCCGCTGCTGCTGCTCCTGGCGGCCCTGGGGACGGGCAGTCCCTGTGTC
TCTGGTGGTTGCCTAAACCTGCAAACATCACCTCTTATCCATCAACATGAAGAATGTCCCTACA
ATGGACTCCACCAGAGGGTCTTCAAGGAGTTAAAGTTACCTACACTGTGCAGTATTTCATCACAA
ATTGGCCCACCAGAGGGTGGCACTGACTACAGATGAGAAGTCCATTCTGTTGTCCTGACAGCTCC
AGAGAAGTGGAAAGAGAAATCCAGAAGACCTTCCTGTTCCATGCAACAAATATACTCCAATCTGA
AGTATAACGTGCTGTGTTGAATACTAAATCAAACAGAACGTGGTCCCAGTGTGACCAACCAC
ACGCTGGTGCTCACCTGGCTGGAGCCGAACACTCTTACTGCGTACACGTGGAGTCCTCGTCCC
AGGGCCCCCTGCCGTGCTAGCCTCTGAGAAGCAGTGTGCCAGGACTTGAAAGATCAATCAT
CAGAGTTCAAGGCTAAAATCATCTTCTGGTATTTGCCATATCTATTACCGTGTGTTCTTT
TCTGTGATGGCTATTCCATCTACCGATATATCCACGTTGGCAAAGAGAAACACCCAGCAAATT
GATTTGATTATGAAATGAATTGACAAAAGATTCTTGTGCCTGCTGAAAAAAATCGTGATTA
ACTTATCACCTCAATATCTGGATGATTCTAAATTCATCAGGATATGAGTTACTGGGA
AAAAGCAGTGATGTATCCAGCCTTAATGATCCTCAGCCAGCGGGAACCTGAGGCCCTCAGGA
GGAAGAGGAGGTGAAACATTAGGTATGCTCGATTGATGGAAATTTTGACTCTGAAG
AAAACACGGAAGGTACTCTCACCCAGCAAGAGTCCCTCAGCAGAACAAACCCCCGGATAAA
ACAGTCATTGAATATGAATATGATGTCAGAACCACTGACATTGTCGGGGCTGAAGAGCAGGA
GCTCAGTTGCAGGAGGGTGTCCACACAAGGAACATTATTGGAGTCGCAGGCAGCGTGGCAG
TCTTGGGCCGAAACGTTACAGTACTCATACACCCCTCAGCTCCAAGACTAGACCCCTGGCG
CAGGAGCACACAGACTCGGAGGAGGGCCGGAGGAAGAGCCATCGACGACCCCTGGCGACTGGGA
TCCCCAAACTGGCAGGCTGTGTATTCTCGCTGTCCAGCTCGACCAGGATTGAGGGCTGCG
AGCCTCTGAGGGGATGGCTCGGAGAGGGTCTCTATCTAGACTCTAGGGAGCCGGCT
CCAGACAGGCCACCAGGAGAAAATGAAACCTATCTCATGCAATTGAGGAATGGGGTTATA
TGTGCAGATGGAAAACTGATGCCAACACTCCTTTGCCTTGTGAAACAAGTGAG
TCACCCCTTGATCCCAGCCATAAGTACCTGGGATGAAAGAAGTTTCCAGTTGTCACTGT
CTGTGAGAATTACTTATTCTCTATTCTCATAGCACGTGTGATTGGTCATGCATGTA
GGTCTCTAACATGATGGTGGCCTCTGGAGTCCAGGGCTGGCCGGTTGTCTATGCAGAGAA
AGCAGTCATAAAATGTTGCCAGACTGGGTGCAGAATTATTACAGGTGGGTGT

FIGURE 76

MSYNGLHQRFKELKLLTLCISISSQIGPPEVALTTDEKSISVVLTAPEKWKRNPEDLPVSMQQIY
SNLKYNVSLNTKSNRTWSQCVTNHTLVLTWLEPNTLYCVHVESFVPGPPRRAQPSEKQCARTLK
DQSSEFKAKIIFWYVLPISITVFLFSVMGYSIYRYIHVGKEKHPANLILYGNEDDKRFFVPAEK
IVINFITLNISDDSKISHQDMSLLGKSSDVSSLNDPQPSGNLRPPQEEEVKHLGYASHLMEIFC
DSEENTEGTSLTQQESLSRTIPPDKTVIEYEYDVRTTDICAGPEEQELSLOEEVSTQGTLLESQA
ALAVLGPQTLQYSYTPQLQDLDPLAQEHTDSEEGPEEPSTTLVDWDPQTGRLCIPSLSSFDQDS
EGCEPSEGDSLGEEGLLSRLYEEPAPDRPPGENETYLMQFMEEWGLYVQMen

Important features:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 140-163

N-glycosylation sites.

amino acids 71-74, 80-83, 89-92, 204-207, 423-426

FIGURE 77

GAGGAGCGGGCCGAGGACTCCAGCGTCCCCAGGTCTGGCATCCTGCACCTGCTGCCCTCTGACAC
CTGGGAAGATGGCCGGCCGTGGACCTCACCCTCTGTGGTTGCTGGCAGCCACCTGATC
CAAGCCACCCCTCAGTCCCCTGCAGTTCTCATCCTCGGCCAAAAGTCATCAAAGAAAAGCTGAC
ACAGGAGCTGAAGGACCACAACGCCACCAGCATCCTGCAGCAGCTGCCGCTGCTCAGTGCATGC
GGGAAAAGCCAGCCGGAGGCATCCCTGTGCTGGCAGCCTGGTGAACACCGCCTGAAGCACATC
ATCTGGCTGAAGGTCACTCACAGCTAACATCCTCAGCTGCAGGTGAAGCCCTGGCCAATGACCA
GGAGCTGCTAGTCAAGATCCCCCTGGACATGGTGGCTGGATTCAACACGCCCTGGTCAAGACCA
TCGTGGAGTCCACATGACGACTGAGGCCAAGCCACCATCCGCATGGACACCCAGTGCAAGTGGC
CCCACCCGCTGGCCTCAGTGAATGTGCCACCAGCCATGGGAGCCTGCGCATCCAATGCTGTA
TAAGCTCTCCTCCTGGTGAACGCCCTAGCTAACAGGTCAATGGCATGTATGCAGAC
CCAATCTAGTGAAAACCAGCTGTGTCCCCTGATCGAGGCTTCCTCAATGGCATGTATGCAGAC
CTCCTGCAGCTGGTGAAGGTGCCATTCCCTCAGCATTGACCGTCTGGAGTTGACCTCTGTA
TCCTGCCATCAAGGGTGAACACCATTCAAGCTCACCTGGGGCCAAGTTGTTGACTCACAGGGAA
AGGTGACCAAGTGGTCAATAACTCTGCAGCTCCCTGACAATGCCAACCTGGACAACATCCCG
TTCAGCCTCATCGTGAGTCAGGACGTGGTGAAGCTGCAGTGGCTGCTGTCTCTCCAGAAGA
ATTCAATGGCCTGTTGACTCTGTGCTCCTGAGAGTGCCATGGCTGAAGTCAAGCATGGC
TGATCAATGAAAAGGCTGCAGATAAGCTGGATCTACCCAGATCGTAAGATCCTAACTCAGGAC
ACTCCGAGTTTTATAGACCAAGGCCATGCCAAGGTGCCCAACTGATCGTGTGGAAAGTGT
TCCCTCCAGTGAAGCCCTCCGCCCTTGTACCCCTGGCATCGAAGCCAGCTCGGAAGCTCAGT
TTTACACCAAAGGTGACCAACTTAACTCAACTGAATAACATCAGCTCTGATCGGATCCAGCTG
ATGAACCTGGGATTGGCTGGTCCAACCTGATGTTCTGAAAAACATCATCACTGAGATCATCCA
CTCCATCCTGCTGCCGAACCAGAAATGGCAAATTAAAGATCTGGGTCCCAGTGTCAATTGGTGAAGG
CCTTGGGATTGAGGCAGCTGAGTCCTCACTGACCAAGGATGCCCTGTGCTTACTCCAGCCTCC
TTGTGGAAACCCAGCTCCTGTCTCCAGTGAAGACTGGATGGCAGCCATCAGGGAAAGGCTGG
GTCCCAGCTGGAGTATGGGTGTGAGCTATAGACCATCCCTCTGCAATCAATAAACACTTG
CCTGTGAAAAA

FIGURE 78

MAGPWTFTLLCGLLAATLIQATLSPTAVLILGPKVIKEKLTQELKDHNATSILQQLPLLSAMREK
PAGGIPVLGSLVNTVLKHIWLVKVTANILQLQVKPSANDQELLVKIPLDMVAGFNTPLVKTIVE
FHMTTEAQATIRMDTSASGPTRLVSDCATSHGSLRIQLLYKLSFLVNALAKQVMNLLVPSLPNL
VKNQLCPVIEASFNGMYADLLQLVKVPISLSIDRLEFDLLYPAIKGDTIQLYLGAKLLDSQGKVT
KWFNNSAASLTMPPTLDNIPFSLIVSQDVVKAAVAAVLSPEEFMVLLDSVLPESAHLKSSIGLIN
EKAADKLGSTQIVKILTQDTPEFFIDQGHAKVAQLIVLEVFPSSEALRPLFTLGIEASSEAQFYT
KGDQLILNLNNISSDRIQLMNSGIGWFQPDVLKNIITEIHSILLPNQNGKLRSGVPVSLVKALG
FEAAESSLTKDALVLTTPASLWKPSPPVSQ

Important features of the protein:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 48-51, 264-267, 401-404

Glycosaminoglycan attachment site.

amino acids 412-415

LBP / BPI / CETP family proteins.

amino acids 407-457

FIGURE 79

GAGAGAAGTCAGCCTGGCAGAGAGACTCTGAAATGAGGGATTAGAGGTGTTCAAGGAGCAAGAGC
TTCAGCCTGAAGACAAGGGAGCAGTCCCTGAAGACGCTTACTGAGAGGTCTGCCATGGCCTCT
CTTGGCCTCCAACTTGTGGCTACATCCTAGGCCTCTGGGGCTTTGGGCACACTGGTGCCT
GCTGCTCCCCAGCTGGAAAACAAGTTCTTATGTCGGTGCAGCATTGTGACAGCAGTTGGCTCT
CCAAGGGCCTCTGGATGGAATGTGCCACACACAGGCACAGGCATCACCCAGTGTGACATCTATAGC
ACCCCTCTGGGCTGCCGCTGACATCCAGGCTGCCAGGCCATGATGGTGACATCCAGTGCAAT
CTCCTCCCTGGCCTGCATTATCTCTGTGGTGGCATGAGATGCACAGTCTCTGCCAGGAATCCC
GAGCAAAGACAGAGTGGCGGTAGCAGGTGGAGTCTTTCATCCTGGAGGCCTCTGGGATTC
ATTCCCTGTTGCCTGGAATCTTCATGGGATCCTACGGGACTTCTACTCACCACGGTGCCTGACAG
CATGAAATTGAGATTGGAGAGGGCTTTACTTGGCATTATTCTCCCTGTTCTCCCTGATAG
CTGGAATCATCCTCTGCTTTCTGCTCATCCCAGAGAAATCGCTCCAACACTACGATGCC
CAAGCCCCAACCTCTGCCACAAGGAGCTCTCCAAGGCCTGGTCAACCTCCAAAGTCAAGAGTGA
GTTCAATTCCCTACAGCCTGACAGGGTATGTGTAAGAACCAGGGCCAGAGCTGGGGGTGGCTG
GGTCTGTGAAAACAGTGGACAGCACCCCGAGGGCCACAGGTGAGGGACACTACCAACTGGATCGT
GTCAGAAGGTGCTGCTGAGGATAGACTGACTTGGCCATTGGATTGAGCAAAGGCAGAAATGGGG
GCTAGTGTAAACAGCATGCAGGTTGAATTGCCAAGGATGCTGCCATGCCAGCCTTCTGTTCC
TCACCTTGCTGCCCTGCCCTAAGTCCCCAACCTCAACTGAAACCCATTCCCTTAAGCCA
GGACTCAGAGGATCCCTTGCCCTCTGGTTACCTGGACTCCATCCCCAAACCCACTAATACA
TCCCAC TGACTGACCCCTGTGATCAAAGACCCCTCTGGCTGAGGTTGGCTCTAGCTCATT
GCTGGGGATGGGAAGGAGAAGCAGTGGCTTTGTGGCATTGCTCTAACCTACTTCTCAAGCTTC
CCTCCAAAGAAACTGATTGCCCTGGAACCTCCATCCCACCTTGTATGACTCCACAGTGTCCA
GACTAATTGTGCATGAACTGAAATAAAACCATCCTACGGTATCCAGGGAACAGAAAGCAGGATG
CAGGATGGGAGGACAGGAAGGCAGCCTGGACATTAAAAAAATA

FIGURE 80

MASLGLQLVGYILGLLGLLGTIVAMLLPSWKTSSYVGASIVTAVGFSKGLWMECATHSTGITQCD
IYSTLLGLPADIQAAQAMMVTTSSAISSLACIISVVGMRCTVFCQESRAKDRVAVAGGVFFILGGL
LGFIPVAWNLHGILRDFYSPLVPDSMKFEIGEALYLGISSLFSLIAGIILCFSCSSQRNRSNYY
DAYQAQPLATRSSPRPGQPPKVKEFNSYSLTGYV

Important features of the protein:

Signal peptide:

amino acids 1-24

Transmembrane domains:

amino acids 82-102, 117-140, 163-182

N-glycosylation site.

amino acids 190-193

PMP-22 / EMP / MP20 family proteins.

amino acids 46-59

FIGURE 81

CCACACCGTCCGCGCCTCTCCCTCTGCTGGACCTTCCTCGTCTCCATCTCTCCCTCCTTC
CCCGCGTTCTCTTCCACCTTCTCTTCCACCTAGACCTCCCTCCCTGCCCTCCTTCCT
GCCCACCGCTGCTTCCCTGCCCTCTCCGACCCCGCTCTAGCAGCAGACCTCCCTGGGGTCTGTGG
GTTGATCTGTGGCCCTGTGCCCTCCGTGTCCTTCGTCTCCCTCCGACTCCGCTCCGG
ACCAGCGGCCTGACCTGGGGAAAGGAGGTCCCGAGGTGAGGGCCTCTCCTGCTGGGA
CTCGCGCTGCTCTGGTCCCCCTGGACTCCCACGCTCGAGCCGCCAGACATGTTCTGCCTTT
CCATGGGAAGAGATACTCCCCGGCGAGAGCTGGCACCCCTACTTGGAGCCACAAGGCCTGATGT
ACTGCCTGCGCTGTACCTGCTCAGAGGGGCCCATGTGAGTTACCGCCTCCACTGTCCGCCT
GTCCACTGCCCTGACGGAGCCACAGCAATGCTGTCCCAAGTGTGTGGAACCTCACAC
TCCCTCTGGACTCCGGCCACCAAAGTCCTGCCAGCACACGGGACCATGTACCAACACGGAG
AGATCTTCAGTGCCCATGAGCTGTTCCCTCCGCCCTGCCAACCAGTGTGTCCCTGCAGCTGC
ACAGAGGGCCAGATCTACTGCCCTCACAACCTGCCCGAACCAAGGCTGCCAGCACCCCTCCC
ACTGCCAGACTCCTGCTGCCAAGCCTGCAAAGATGAGGCAAGTGAGCAATCGGATGAAGAGGACA
GTGTGCAGTCGCTCCATGGGGTGGAGACATCCTCAGGATCCATGTTCCAGTGTGCTGGAGAAAG
AGAGGCCGGGCACCCAGCCCCACTGGCCTCAGGCCCTCTGAGCTTCATCCCTGCCACTT
CAGACCCAAGGGAGCAGGCAGCACAACTGTCAAGATCGTCTGAAGGAGAACATAAGAAAGCCT
GTGTGCATGGCGGGAGACGTACTCCCACGGGAGGTGTGGCACCCGGCTTCCGTGCCCTCGC
CCCTGCCCTGCATCCTATGCACCTGTGAGGATGCCGCCAGGACTGCCAGCGTGTGACCTGTCC
CACCGAGTACCCCTGCCGTACCCCGAGAAAGTGGCTGGGAAGTGTGCAAGATTGCCAGAGG
ACAAAGCAGACCCCTGCCACAGTGAGATCAGTTCTACCAGGTGTCCCAAGGCACCGGGCGGGTC
CTCGTCCACACATCGGTATCCCCAAGCCCAGACAAACCTGCGTCGCTTGCCTGGAACACGAGGC
CTCGGACTTGGTGGAGATCTACCTCTGGAAGCTGGTAAAGATGAGGAAACTGAGGCTCAGAGAG
GTGAAGTACCTGCCCAAGGCCACACAGCCAGAATCTTCACTTGACTCAGATCAAGAAAGTCAG
GAAGCAAGACTCCAGAAAGAGGCACAGCACTCCGACTGCTCGCTGGCCCCACGAAGGTCACT
GGAACGTCTCCTAGCCCAGACCCCTGGAGCTGAAGGTACGCCAGTCCAGACAAAGTACCAAG
ACATAACAAAGACCTAACAGTTGCAGATATGAGCTGTATAATTGTTATTATATATTAATAAA
TAAGAAGTTGCATTACCCCTAAAAAAAAAAAAAAAAAAAAA

FIGURE 82

MVPEVRVLSSLLGLALLWFPLDSHARARPDMFCLFHGKRYSPGESWHPYLEPQGLMYCLRCTCSE
GAHVSCYRLHCPPVHCPQPVTEPQQCCPKCVEPHTPSGLRAPPKSCQHNGTMYQHGEIFSAHELF
PSRLPNQCVLCSCTEGQIYCGLTTCPEPGCPAPLPLPDSCCQACKDEASEQSDEEDSVQSLHGVR
HPQDPCSSDAGRKRGPGTPAPTGLSAPLSFIPRHFRPKGAGSTTVKIVLKEHKKACVHGGKTYs
HGEVWHPAFRAFGPLPCILCTCEDGRQDCQRVTCPTEYPCRHPEKVAGKCKICPEDKADPGHSE
ISSTRCPKAPGRVLVHTSVSPSPDNLRRFALEHEASDLVEIYLWKLVKDEETEAQRGEVPGPRPH
SQNLPLSDQESQEARNPERGTALPTARWPPRSLERLPSDPPGAEGHGQSRQSDQDITKT

Signal peptide:

amino acids 1-25

FIGURE 8.3

GACAGCTGTCTCGATGGAGTAGACTCTAGAACAGCGCAGTTGCCCTCCGCTCACGCAGAGCCTCTCC
GTGGCTTCCGCACCTTGAGCATTAGGCCAGTTCTCCCTCTCTAATCCATCCGTACACCTCTCCTGTCA
TCCGTTCCATGCCGTGAGGTCCATTACAGAACACATCCATGGCTCTCATGCTCAGTTGGTCTGAGTC
TCCTCAAGCTGGGATCAGGGCAGTGGCAGGTGTTGGGCCAGACAAAGCCTGTCCAGGCCTTGGTGGGGAG
GACGCAGCATTCTCCTGTTCTGTCTCCTAAGACCAATGCAGAGGCCATGGAAGTGCAGGTTCTCAGGGG
CCAGTTCTCTAGCGTGGTCCACCTCTACAGGGACGGGAAGGACCAGCCATTATGCAGATGCCACAGTATC
AAGGCAGGACAAAATGGTGAAGGATTCTATTGCGGAGGGCGCATCTCTGAGGCTGGAAAACATTACT
GTGTTGGATGCTGGCCTCTATGGGTGCAGGATTAGTCCCAGTCTTACTACCAGAACGCCATCTGGAGCT
ACAGGTGTCAGCACTGGCTCAGTTCTCATTCCATCACGGATATGTTGATAGAGACATCCAGCTAC
TCTGTCAGTCCTCGGGCTGGTCCCCGGCCACAGCGAAGTGGAAAGGTCCACAAGGACAGGATTGTCC
ACAGACTCCAGGACAAACAGAGACATGCATGGCTGTTGATGTGGAGATCTCTGACCGTCCAAGAGAA
CGCCGGAGCATATCCTGTTCCATGCCATGCTCATCTGAGCCGAGAGGTGGAATCCAGGGTACAGATAG
GAGATACCTTTTCGAGCCTATATCGTGGCACCTGGCTACCAAAGTACTGGGAATACTCTGCTGTGGCCTA
TTTTTGGCATTGTTGACTGAAGATTCTTCTCAAATTCCAGTGGAAAATCCAGGCGGAACGGACTG
GAGAAGAAAGCACGGACAGGCAGAATTGAGAGACGCCGGAAACACGCAGTGGAGGTGACTCTGGATCCAG
AGACGGCTCACCGAAGCTCTCGCTTCTGATCTGAAAACGTGTAACCCATAGAAAAGCTCCCCAGGAGGTG
CCTCACTCTGAGAAGAGATTACAAGGAAGAGTGTGGTGGCTCTCAGAGTTCCAAGCAGGGAAACATTA
CTGGGAGGTGGACGGAGGACACAATAAAGGTGGCGCGTGGAGTGTGCCGGATGATGTGGACAGGAGGA
AGGAGTACGTACTTGTCTCCGATCATGGTACTGGCTCTCAGACTGAATGGAGAACATTGTATTTC
ACATTAATCCCCGTTTATCAGCGTCTTCCCAGGACCCCACCTACAAAATAGGGTCTCCTGGACTA
TGAGTGTGGACCACATCCTCTCAACATAATGACCAGTCCTTATTATACCCGACATGTCGTTG
AAGGCTTATTGAGGCCCTACATTGAGTATCCGTCTATAATGAGCAAATGGAACCTCCATAGTCATCTGC
CCAGTCACCCAGGAATCAGAGAAAGAGGCCCTTGGCAAAGGGCTCTGCAATCCCAGAGACAAGAACAG
TGAGTCCTCCTCACAGGCAACCACGCCCTCCTCCCCAGGGTGAAATGTAGGATGAATCACATCCCACAT
TCTTCTTTAGGGATATTAGGTCTCTCCCAGATCCAAGTCCCGCAGCAGCCGCCAAGGTGGCTTCCA
GATGAAGGGGACTGGCCTGTCACATGGAGTCAGGTGTCAGGCTGCCCTGAGCTGGAGGGAAAGAAG
CTGACATTACATTAGTTGCTCTCACTCCATCTGGCTAAGTGAATCTTGAATACCAACCTCTCAGGTGAAG
AACCGTCAGGAATTCCCATCTCACAGGCTGTGGTAGATTAAGTAGACAAGGAATGTGAATAATGCTTAG
ATCTTATTGATGACAGAGTGTATCCTAATGGTTGTCATTATATTACACTTCAGTAAAAAA

FIGURE 84

MALM LSLV SLLKL GSG QWQVFG PDKPV QALV GEDA AFSC FLSP KTNA EAME VRFF RGQF SS VVH
LYRD GKDQ PFMQ MPQY QGRT KLV KDSIA EGRIS LRL ENIT VLDAG LYGC RISSQ SYYQ KAI WELQ
VSAL GSVPL ISIT GYVDRD IQLLCQSSGWF PRPTAKWKG P QGQDLSTD SRTN RDMH GLFD VEISL
TVQENAGSISCSMRHAHLSREVESRVQIGDTFFEPISWHLATKVLGILCCGLFFGIVGLKIFFSK
FQWKIQAELDWRRKHGQAE LRDAK HAVEV TLD PETA HPKLCVSDLKTVTHRKA P QEVPHSEKRF
TRKSVVASQSFQAGKHYWEVDGGHNKRWRVGVCRDDVDRRKEYVTLSPDHGYWVLRLNGEHLYFT
LNPRFISVFPRT PPTKIGVFLDYECGTISFFNINDQSLIYTLTCRFEGLLRPYIEYPSYNEQNGT
PIVICPVTQESEKEASWQRASAIPETSNSESSQATT PFLPRGEM

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 239-255

FIGURE 85

AACAGACGTTCCCTCGCGGCCCTGGCACCTTAACCCCAGACATGCTGCTGCTGCTGCCCT
GCTCTGGGGAGGGAGAGGGCGGAAGGACAGACAAGTAAACTGCTGACGATGCAGAGTTCCGTGA
CGGTGCAGGAAGGCCTGTGTCCATGTGCCCTGCTCCTCTCCTACCCCTCGCATGGCTGGATT
TACCCCTGGCCCAGTAGTTCATGGCTACTGGTCCGGAAAGGGGCAATACAGACCAGGATGCTCC
AGTGGCCACAAACAACCCAGCTGGGAGTGTGGGAGGAGACTGGGACCGATTCCACCTCCTG
GGGACCCACATACCAAGAATTGCACCCCTGAGCATTCAGAGATGCCAGAAGAAGTGATGATGCCGGGAGA
TACTTCTTCGTATGGAGAAAGGAAGTATAAAATGGAATTATAAACATCACCGGCTCTGTGAA
TGTGACAGCCTTGACCCACAGGCCAACATCCTCATCCCAGGCACCCCTGGAGTCCGGCTGCC
AGAATCTGACCTGCTCTGTGCCCTGGCCTGTGAGCAGGGACACCCCTATGATCTCCTGGATA
GGGACCTCCGTGTCCCCCTGGACCCCTCCACCACCCGCTCCTCGGTGCTCACCTCATCCCACA
GCCCCAGGACCATGGCACCAGCCTCACCTGTCAGGTGACCTCCCTGGGCCAGCGTGACCACGA
ACAAGACCGTCCATCTCAACGTGTCTACCCGCTCAGAACTGACCATGACTGTCTCCAAGGA
GACGGCACAGTATCCACAGTCTGGAAATGGCTCATCTGTCACTCCAGAGGGCAGTCT
GCGCCTGGTCTGTGCAGTGCAGTGCAGCAGCATCCCCCTGCCCAGGCTGAGCTGG
GAGGCCTGACCCGTGCCCCTCACAGCCTCAAACCCGGGGTGCTGGAGGCTGCCTGGGTGCAC
CTGAGGGATGCAGCTGAATTACCTGCAGAGCTCAGAACCCCTTCGGGTCTCAGCAGGTCTACCT
GAACGTCTCCCTGCAGAGCAAAGCCACATCAGGAGTGACTCAGGGGTGGTGGGGAGGCTGGG
CCACAGCCCTGGTCTCCCTGTCCTTCTGCGTCATTCGTGTAGTGAGGTCTGCAGGAAGAAA
TCGGCAAGGCCAGCGCCGGGTGGGAAGGATACGGGCATAGAGGATGCAAACGTGTCAGGGGTTC
AGCCTCTCAGGGGCCCTGTACTGAACCTGGGCAGAAAGACAGTCCCCCCAGGACCCACCCCTCCCCCA
CTTCTGCCCGCTCTCAGTGGGGGAAGGAAGAGCTCCAGGTATGCATCCCTCAGGCTCCAGATGGT
AAGCCTGGGACTCGGGGGACAGGGAGCCACTGAACCCGAGTACTCGGAGAATCAAGATCCACAG
ATGAGAAACTGCAGAGACTCACCCTGATTGAGGGATCAGCCCCTCCAGGAAGGGAGGAAGGTCA
GAGGGTATCTGTAGATTAACAGCCCCTCAACGTGATGAGCTATGAAACACTATGATTATGA
TGCAGAGTGAAAAGCACACAGGCTTAGAAGGTCAAAAGATCTCAAAACCTGAATCCACACGTGCCC
TCCCTTTTTTTTTAACTAAAGACAGACAAATCCCTTA

FIGURE 86

MLLPLLWGRERAEGQTSKLLTMQSSVTQEGLCVHPCFSYPSHGWIYPGPVHGYWFREG
ANTDQDAPVATNNPARAVWEETRDRFHLLGDPHTKNCTL SIRDARRSDAGRYFFRMEKGSIKWNY
KHHRLSVNVTALTHRPNILIPGTLESGCPQNLTCSPWACEQGTPPMISWIGTSVSPLDPSTTRS
SVLTLIPQPQDHGTSLTCQVTFP GASVTTNKT VHLN VS YPPQNL TMTV FQGDGT VSTVLNGSSL
SLPEGQSLRLVCAVDAVDSNPPARLSLSWRGLTLCPSPNPGVLELPWVHLRDAAEFTCRAQNP
LGSQQVYLNVS LQSKATSGVTQGVVGGAGATALVFLSFCVIFVVVRSCRKKSARPAAGVGDTGIE
DANAVRGSASQGPLTEPWAEDSPPDQPPPASARSSVGEGELOQYASLSFQMVKPWDSRGQEATDTE
YSEIKIHR

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 351-370

FIGURE 87

AGAAAGCTGCACTCTGTTGAGCTCCAGGGCGCAGTGGAGGGAGGTGAAGGAGCTCTGTAC
CCAAGGAAAGTGCAGCTGAGACTCAGACAAGATTACAATGAACCAACTCAGCTTCTGCTGTT
TCATAGCGACCACCAGAGGATGGAGTACAGATGAGGCTAATACTTACTTCAAGGAATGGACCTGT
TCTTCGTCTCCATCTGCCAGAAGCTGCAAGGAAATCAAAGACGAATGTCTAGTGCATTGA
TGGCCTGTATTTCTCCGACTGAGAATGGTGTATCTACCAGACCTCTGTGACATGACCTCTG
GGGGTGGCGGCTGGACCCTGGTGGCCAGCGTGCATGAGAATGACATGCGTGGAAAGTGCACGGT
GGCGATCGCTGGTCCAGTCAGCAGGGCAGCAAAGCAGACTACCCAGAGGGGACGGCAACTGGGC
CAACTACAACACCTTGGATCTGCAGAGGCAGGCCACGAGCGATGACTACAAGAACCTGGCTACT
ACGACATCCAGGCCAAGGACCTGGCATCTGGCACGTGCCAATAAGTCCCCATGCAGCACTGG
AGAAACAGCTCCCTGCTGAGGTACCGCACGGACACTGGCTCCTCAGACACTGGACATAATCT
GTTTGGCATCTACCAGAAATATCCAGTGAATATGGAGAAGGAAAGTGGACTGACAACGGCC
CGGTGATCCCTGTGGTCTATGATTTGGCGACGCCAGAAAACAGCATCTTATTACTCACCCTAT
GGCCAGCGGAATTCACTGCAGGGATTTGTCAGTCAGGGTATTAATAACGAGAGAGCAGCCAA
CGCCTTGTGTGCTGGAATGAGGGTCACCGATGTAACACTGAGCATCACTGCATTGGAGGAG
GATACTTTCCAGAGGCCAGTCCCCAGCAGTGTGGAGATTCTGGTTGATTGGAGTGGATAT
GGAACTCATGTTGGTTACAGCAGCAGCGTGGAGATAACTGAGGCAGCTGTGCTCTATTCTATCG
TTGAGAGTTGTGGAGGGAACCCAGACCTCTCCTCCACCAGAGATCCAAAGGATGGAGAA
CAACTTACCCAGTAGCTAGAATGTTAATGGCAGAAGAGAAAACAATAATGACTCAAGA
AAAAAA

FIGURE 88

MNQLSFLLFLIATTRGWSTDEANTYFKEWTCSSPSLPRSCKEIKDECPSAFDGLYFLRTENGVI
YQTFCDMTSGGGGWTLVASVHENDMRGKCTVGDRSSQQGSKADYPEGDGNWANYNTFGSAEAAT
SDDYKNPGYYDIQAKDLGIWHVPNKS PMQHWRNSSLRYRTDTGFLQTLGHNLFGIYQKYPVKYG
EGKCWTDNGPVI PVVYDFGDAQKTASYYSPYGQREFTAGFVQFRVFNNEAANALCAGMRVTGCN
TEHHCIGGGGYFPEASPQQCGDFSGFDWSGYGTHVGYSSSREITEAAVLLFYR

Important features:

Signal peptide:

amino acids 1-16

N-glycosylation site.

amino acids 163-167

Glycosaminoglycan attachment sites.

amino acids 74-78, 289-293

N-myristoylation sites.

amino acids 76-82, 115-121, 124-130, 253-259, 292-298

FIGURE 89

CTAGATTGTCGGCTGCGGGGAGACTTCAGGAGTCGCTGTCTGAACCTCCAGCCTCAGAGAC
CGCCGCCCTGTCCCCGAGGGCCATGGGCCGGTCTCAGGGCTTGTGCCCTCTCGCTTCTGACG
CTCCTGGCGCATCTGGTGGTCGTACACCTTATTCTGGTCCCAGGACAGAACATACAGGCCTG
CCTGCCTCTCACGTTACCCCCGAGGAGTATGACAAGCAGGACATTAGCTGGTGGCCGCGCTCT
CTGTCACCCCTGGGCCTCTTGAGTGGAGCTGGCCGGTTCCCTCTCAGGAGTCTCCATGTTAAC
AGCACCCAGAGCCTCATCTCCATTGGGCTCACTGTAGTGCATCCGTGGCCCTGTCCTTCTTCAT
ATTGAGCGTTGGGAGTGCACGTACGTATTGGTACATTTGTCTTCTGCAGTGCCTTCCAGCTG
TCACTGAAATGGCTTATTGTCACCGTCTTGGCTGAAAAAGAAACCTCTGATTACCTCA
TGACGGAACCTAAGGACGAAGCCTACAGGGCAAGGGCCGCTCGTATTCTGGAAAGAAGGAAG
GCATAGGCTTCGGTTTCCCCTCGAAACTGCTCTGCTGGAGGATATGTGTTGGAAATAATTACG
TCTTGAGTCTGGATTATCCGCATTGTATTTAGTGCTTGTAAATAAAATATGTTTAGTAACA
TTAAGACTTATACAGTTAGGGACAATTAAAAAAAAAAAAA

FIGURE 90

MGRVSGLVPSRFLTLLAHLVVVITLFWSRDSNIQACLPLTFTPPEEYDKQDIQLVAALSVTLGLFA
VELAGFLSGVSMFNSTQSLISIGAHCSASVALSFFIFERWECTTYWYIFVFCSALPAVTEMALFV
TVFGLKKKPF

Transmembrane domain:

amino acids 12-28 (type II), 51-66, 107-124

FIGURE 91

CTGGGACCCCGAAAAGAGAAGGGGAGAGCGAGGGGACGAGAGCGGAGGAGGAAGATGCAACTGAC
TCGCTGCTGCTCGTGTTCCTGGTGCAGGGTAGCCTCTATCTGGTCATCTGTGGCCAGGATGATG
GTCCTCCCGGCTCAGAGGACCCCTGAGCGTGTGACCAACGAGGGCCAGCCCCGGCCCCGGTGCCT
CGGAAGCGGGGCCACATCTCACCTAACGTTCCGCCCCATGGCCAATTCCACTCTCCTAGGGCTGCT
GGCCCCGCCTGGGGAGGGCTTGGGCATTCTTGGCAGCCCCCAACCGCCCGAACACAGCCCC
CACCCCTCAGCCAAGGTGAAGAAAATCTTGGCTGGGCAGCTACTCCAACATCAAGACGGTG
GCCCTGAACCTGCTCGTCACAGGGAAAGATTGTGGACCATGGCAATGGACCTTCAGCGTCCACTT
CCAACACAATGCCACAGGCCAGGGAAACATCTCCATCAGCCTCGTGGCCCCAGTAAGCTGTAG
AGTTCCACCAGGAACACAGCAGATCTTCATCGAACGCCAAGGCCTCCAAAATCTCAACTGCCGGATG
GAGTGGGAGGAAGGTAGAACGGGCCGGACCTCGCTTGACCCACGACCCAGCCAAGATCTG
CTCCCGAGACCACGCTCAGAGCTCAGCCACCTGGAGCTGCTCCCAGGCCCTCAAAGTCGTCTGTG
TCTACATGCCCTCTACAGCACGGACTATCGGCTGGTCCAGAAGGTGTGCCAGATTACAACATAC
CATAGTGATAACCCCTACTACCCATCTGGGTGACCCGGGGCAGGCCACAGAGGCCAGGGC
TGGAAGGACAGGCCTGCCATGCAGGAGACCATCTGGACACCAGGGCAGGGAAAGGGTTGGCCTC
AGGCAGGGAGGGGGGGTGGAGACGAGGAGATGCCAAGTGGGCCAGGGCCAAGTCTCAAGTGGCAG
AGAAAGGGTCCAAGTGCTGGTCCAACCTGAAGCTGTGGAGTGACTAGATCACAGGAGCACTGG
AGGAGGAGTGGCTCTGTGCAGCCTCACAGGGCTTGCCACGGGCCACAGAGAGATGCTGG
TCCCCGAGGCCTGTGGCAGGCCGATCAGTGTGGCCCCAGATCAAGTCATGGAGGAAGCTAAGC
CCTTGGTTCTGCCATCCTGAGGAAAGATAGCAACAGGGAGGGAGATTTCATCAGTGTGGACA
GCCTGTCAACTTAGGATGGATGGCTGAGAGGGCTTCCTAGGAGCCAGTCAGCAGGGTGGGTGG
GCCAGAGGAGCTCCAGCCCTGCCTAGTGGGCCCTGAGCCCTGTGCTGAGCATGG
CATGAGGCTGAAGTGGCAACCCTGGGTCTTGATGTCTTGACAGATTGACCATCTGTCTCCAGC
CAGGCCACCCCTTCCAAAATCCCTCTGCCAGTACTCCCCCTGTACCAACCCATTGCTGATG
GCACACCCATCCTAACGTAAGACAGGACGATTGTGGCCTCCCACACTAACGCCACAGCCCAC
CGCGTGTGTGTCCCTTCCACCCCAACCCCTGCTGGCTCTGGAGCATCCATGTCCCG
GAGAGGGTCCCTAACAGTCAGCCTCACCTGTCAAGACCGGGTTCTCCGGATCTGGATGGCG
CGCCCTCTCAGCAGCGGGCACGGTGGGCGGGCGGGCGAGAGCATGTGCTGGATCTGTT
TGTGTGTCTGTCTGTGGTGGGGAGGGAGGGAAAGTCTTGTAAACCGCTGATTGCTGACTTT
TGTGTGAAGAATCGTGTCTGGAGCAGGAAATAAGCTGCCCGGGCA

FIGURE 92

MQLTRCCFVFLVQGSILYLVICQDDGPPGSEDPERDDHEGQPRPRVPRKRGHISPKSRPMANSTL
LGLLAPPGEAWGILGQPPNRPNHSPPPSAKVKKIFGWGDFYSNIKTVALNLLVTGKIVDHGNGTF
SVHFQHNATGQGNISISLVPPSKAVEFHQEQQIFIEAKASKIFNCRMWEKVERGRRTSLCTHDP
AKICSRDHAQSSATWSCSQPKVVCVYIAFYSTDYRLVQKCPDNYHSDTPYYPSG

Important features of the protein:

Signal peptide:

amino acids 1-14

N-glycosylation sites.

amino acids 62-65, 127-130, 137-140, 143-146

2-oxo acid dehydrogenases acyltransferase

amino acids 61-71

FIGURE 93

CGGTGGCC**ATG**ACTGCGGCCGTGTTCTCGGCTGCGCCTCATTGCCCTCGGGCCTGCGCTGCC
CTTTATGTCTCACCATGCCATCGAGCCGTTGCGTATCATCTCCTCATGCCGGAGCTTCTT
CTGGTTGGTGTCTACTGATTGTCGCTTGTGGTTCATGGCAAGAGTCATTATTGACAACA
AAGATGGACCAACACAGAAATATCTGCTGATCTTGGAGCGTTGCTCTGCTATATCCAAGAA
ATGTTCCGATTGCATATTATAAACTCTTAAAAAGCCAGTGAAGGTTGAAGAGTATAAACCC
AGGTGAGACAGCACCCCTATGCGACTGCTGCCATGTTCTGGCTGGGCTTGAATCATGA
GTGGAGTATTTCCTTGTGAATAACCTATCTGACTCCTGGGCCAGGCACAGTGGCATTCAT
GGAGATTCTCCTCAATTCTCCTTATTGACTGCTGGCTGACGCTGGTCATTATCTGCTGCATGT
ATTCTGGGCATTGTATTTTGATGGCTGTGAGAAGAAAAAGTGGGCATCCTCCTATCGTC
TCCTGACCCACCTGCTGGTGTAGCCCAGACCTCATAAGTTCTTATTATGGAATAAACCTGGCG
TCAGCATTATAATCCTGGTGCTCATGGCACCTGGCATTCTAGCTGCAGGAGGAGCTGCCG
AAGCCTGAAACTCTGCCTGCTGCCAAGACAAGAACTTCTTACAACCAGCGCTCCAGAT
AACCTCAGGGAACCGACCTCCAAACCGCAGACTACATCTTAGAGGAAGCACAAC**GTGCCT**
TTTCTGAAAATCCCTTTCTGGTGGATTGAGAAAGAAATAAAACTATGCAGATA

FIGURE 94

MTAAVFFGCAFIAFGPALALYVFTIAIEPLRIIFLIAGAFFWLVSLLISSLVWFMARVIIDNKDG
PTQKYLIFGAFVSVYIQEMFRFAYYKLLKKASEGLKSINPGETAPSMRLLAYVSGLGFGIMSGV
FSFVNTLSDSLGP GTVGIHGDSPQFFLYSAFMTLVIILLHVFWGIVFFDGCEKKW GILLIVLLT
HLLVSAQTFIISYYGINLASAFIILVLMGTWAFLAAGGSCRSLKLCLLCQDKNFLLYNQSR

Important features of the protein:

Signal peptide:

amino acids 1-19

Transmembrane domains:

amino acids 32-51, 119-138, 152-169, 216-235

Glycosaminoglycan attachment site.

amino acids 120-123

Sodium:neurotransmitter symporter family protein

amino acids 31-65

FIGURE 95

AATTTTCACCAGAGTAAACTGAGAAACCAACTGGACCTTGAGTATTGTACATTTGCCTCGTG
GACCCAAAGGTAGCAATCTGAAACATGAGGAGTACGATTCTACTGTTTGCTTCTAGGATCAAC
TCGGTCATTACCACAGCTAAACCTGCTTGGACTCCCTCCCACAAAACTGGCTCCGGATCAGG
GAACACTACCAAACCAACAGCAGTCAAATCAGGTCTTCCTTAAGTCTGATACCATTAAACA
CAGATGCTCACACTGGGCCAGATCTGCATCTGTTAAATCCTGCTGCAGGAATGACACCTGGTAC
CCAGACCCACCCATTGACCCTGGGAGGGTGAATGTACAACAGCAACTGCACCCACATGTGTTAC
CAATTTGTCACACAACTTGGAGCCCAGGGCACTATCCTAAGCTCAGAGGAATTGCCACAAATC
TTCACGAGCCTCATCCATTCTGTTCCGGAGGCATCCTGCCACCAGTCAGGCAGGGC
TAATCCAGATGTCCAGGATGGAAGCCTCCAGCAGGAGGAGCAGGTGTAAATCCTGCCACCCAGG
GAACCCCAGCAGGCCCTCCAACTCCCAGTGGCACAGATGACGACTTGCAGTGACCACCCCT
GCAGGCATCCAAAGGAGCACACATGCCATCGAGGAAGCCACCACAGAATCAGCAAATGGAATTCA
GTAAGCTTTCAATTTCAACTAAGCTGCCTCGAATTGGTGATACTGTGAATCTTATC
ATTGATTATATTATGGAATAGATTGAGACACATTGGATAGTCTTAGAAGAAATTAAATTCTTAATT
TACCTGAAATATTCTGAAATTTCAGAAATATGTTCTATGTAGAGAATCCAACTTTAAAAAA
CAATAATTCAATGGATAAATCTGTCTTGAAATATAACATTATGCTGCCTGGATGATGCATAT
TAAAACATATTGGAAACTGGAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 96

MRSTILLFCILLGSTRSLPQLKPALGLPPTKLAPDQGTLPNQQQSNQVFPSSLIPLTQM
LTLPDLHLLNPAAGMTPGTQTHPLTLGGLNVQQQLHPHVLPPIFVTQLGAQGTILSSEE
LPQIFTSLIIHSLFPGGILPTSQAGANPDVQDGSLPAGGAGVNPATQGTPAGRLPTPSG
TDDDFAVTTPAGIQRSTHAIEEATTESANGIQ

Signal peptide:

amino acids 1-16

FIGURE 97

GCTCAAGTGCCTGCCTGCCCAACCCAGCCCAGCCTGGCCAGAGCCCCCTGGAGAAGGAGCTCT
CTTCTGCTGGCAGCTGGACCAAGGGAGCCAGTCTGGGCGCTGGAGGGCCTGTCCTGACCATG
GTCCTGCCTGGCTGTGGCTGTTGTCTCCGTCCCCCCAGGCTCTCCCAAGGCCAGCCTGC
AGAGCTGTCTGTGGAAGTCCAGAAAATCTGGTGGAAATTCCCTTATACCTGACCAAGTTGC
CGCTGCCCGTGAAGGGGCTGAAGGCCAGATCGTGTGTCAGGGACTCAGGCAAGGCAACTGAG
GGCCCATTTGCTATGGATCCAGATTCTGGCTTCCTGCTGGTACCCAGGGCCTGGACCGAGAGGA
GCAGGCAGAGTACCAAGTACAGGTACCCCTGGAGATGCAGGATGGACATGTCTGTGGGTCCAC
AGCCTGTGCTTGTGCACGTGAAGGATGAGAATGACCAGGTGCCCATTTCTCAAGCCATCTAC
AGAGCTCGGCTGAGCCGGGTACCCAGGCTGGCATCCCCCTTCCTGAGGCTTCAGACCG
GGATGAGCCAGGCACAGCCAACTCGGATCTCGATTCCACATCCTGAGCCAGGCTCCAGCCCAGC
CTTCCCCAGACATGTTCCAGCTGGAGCCTGGCTGGCTGGGGCTCTGGCCCTCAGCCCCAAGGGGAGC
ACCAGCCTTGACCAACGCCCTGGAGAGGACCTACAGCTGGTACAGGTCAAGGACATGGGTGA
CCAGGCCTCAGGCCACCAGGCCACTGCCACCGTGAAGTCTCCATCATAGAGAGCACCTGGGTGT
CCCTAGAGCCTATCCACCTGGCAGAGAATCTCAAAGTCCTATACCGCACCACATGGCCAGGTA
CACTGGAGTGGGGGTATGTGACTATCACCTGGAGAGGCCATCCCCGGGACCCCTTGAAGTGA
TGCAGAGGAAACCTTACGTGACCAAGAGCTGGACAGAGAAGCCCAGGCTGAGTACCTGCTCC
AGGTGCGGGCTCAGAATTCCATGGCAGGACTATGCGGCCCCCTCTGGAGCTGCACGTGCTGGT
ATGGATGAGAATGACAACGTGCCTATGCCCTCCCCGTGACCCACAGTCAGCATCCCTGAGCT
CAGTCCACCAAGGTACTGAAGTGAATGACTAGACTGTCAGCAGAGGATGCAGATGCCCGGCTCCCCA
ATTCCACGTTGTATCAGCTCTGAGCCCTGAGCCTGAGGATGGGTAGAGGGAGAGCCTTC
CAGGTGGACCCCACTCAGGCAGTGTGACGCTGGGGTGCCTCCACTCCGAGCAGGCCAGAACAT
CCTGCTTCTGGTGCCTGGCATGGACCTGGCAGGGCAGAGGGTGGCTCAGCAGCACGTGTGAAG
TCGAAGTCGCACTGAAGATATCAATGATCACGCCCTGAGTTCATCACTTCCAGATTGGCCT
ATAAGCCTCCCTGAGGATGTGGAGCCGGACTCTGGTGGCCATGCTAACAGCATTGATGCTGA
CCTCGAGCCCGCTTCCGCCTATGGATTTCGCATTGAGAGGGAGACACAGAACGGACTTTG
GCCTGGATTGGGAGCCAGACTCTGGCATGTTAGACTCAGACTCTGCAAGAACCTCAGTTATGAG
GCAGCTCCAAGTCATGAGGTGGTGGTGGTGCAGAGTGTGGCAAGCTGGTGGGGCAGGCC
AGGCCCTGGAGCCACCGCACGGTACGTGCTAGTGGAGAGAGTGTGACAGGCCACCCAAAGTTGG
ACCAGGAGAGCTACGAGGCCAGTGTCCCCATCAGTGCCTGCCAGCCGGCTTTCTGTCGACCATC
CAGCCCTCCGACCCCATCAGCCGAACCTCAGGTTCTCCCTAGTCATGACTCAGAGGGCTGGCT
CTGCATTGAGAAATTCTCCGGGGAGGTGCACACCGCCAGTCCCTGGTCCAAACCCACGGTGCAACGGGATTGG
ACACCTACACGGTGTGGAGGCCAGGATACAGCCCTGACTCTTGCCTGTCGCCCCCTCCAA
TACCTCTGCACACCCGCCAAGACCATGGCTTGATCGTAGTGGACCCAGCAAGGACCCGATCT
GGCCAGTGGCACGGTCCCTACAGCTCACCCCTGGTCCAAACCCACGGTGCAACGGGATTGG
GCCTCCAGACTCTCAATGGTCCCATGCCTACCTCACCTGGCCCTGCATTGGTGGAGGCCACGT
GAACACATAATCCCCGTGGTGGTCAGCCACAATGCCAGATGTGGCAGCTCTGGTCAGTGAT
CGTGTGCGTGCACAGTGGAGGGCAGTGCATGCGCAAGGTGGCCGATGAAGGGCATGCCA
CGAAGCTGTCGGCAGTGGCATTCTGTAGGCACCCCTGGTAGCAATAGGAATCTCCTCATCCTC
ATTTCAACCACTGGACCATGTCAAGGAAGAAGGGACCCGGATCAACCAGCAGACAGCGTGCCT
GAAGGCGACTGTCTGAATGGCCAGGCAGCTAGCTGGAGCTGGCTCTGGCTCCATCTGAG
TCCCTGGGAGAGAGGCCAGCACCCAAAGATCCAGCAGGGACAGGACAGAGTAGAAGCCCCCTCCA
TCTGCCCTGGGGTGGAGGCACCATCACCATCACAGGATGTCTGCAGAGCCTGGACACCAACTT
TATGGACTGCCCATGGAGTGTCAAATGTCAGGGTGTGCCCCAATAATAAGCCCCAGAGAA
CTGGGCTGGCCCTATGGAAAAAAGAAAAAAGAAAAAAGAAAAAAG

FIGURE 98

MVPAWLWLLCVSVPQALPKAQPAELSVEPVENYGGNFPLYLTKLPLPREGAEGQIVLSGDSGKAT
EGPFAMDPDSGFLLVTRALDREEQAEYQLQVTLEMQDGHVLWGPQPVLVHKDENDQVPHFSQAI
YRARLSRGTRPGIPFLFLEASDRDEPGTANSDLRFHILSQAPAQPSPDMFQLEPRLGALALSPKG
STS LDHALERTYQLLVQVKDMGDQASGHQATATVEVSIIESTWVSLPEIHLAENLKVLYPHHMAQ
VHWSSGGDVHYHLESHPPGPFEVNAEGNLYVTRELDREAQAEYLLQVRAQNSHGEDYAAPPLELHVL
VMDENDNVPICPPRDPTVSIPELSPPGTEVTRLSAEDADAPGSPNSHVYQLLSPEPEDGVEGRA
FQVDPTSGSVTLGVLPLRAGQNILLVLAMDLAGAEGGFSSCTCEVEVAVTDINDHAPEFITSQIG
PISLPEDVEPGTLVAMLTAIDADLEPAFRLMDFAIERGDTEGTFG LDWEPSGHVRLRLCKNLSY
EAAPSHEVVVVVQSVAKLVGPGPGATATVTVLVERVMPPPQLDQESYEASVPISAPAGSFLLT
IQPSDPISRTLRFSLVNDSEGWL CIEKFSGEVHTAQSLQGAQPGDTYTVLVEAQDTALTLPVPS
QYLCTPRQDHGLIVSGPSKDPDLASGHGPYSFTLGPNPTVQRDWRLQTLNGSHAYLTLALHWEP
REHIIIPVVVSHNAQM WQLLVRVIVCRCNVEGQCMRKVGRMKGMPTKLSAVGILVGTLVAIGIFLI
LIFTHWTMSRKKDQPADSVPLKATV

Signal peptide:

amino acids 1-18

Transmembrane domain:

amino acids 762-784

FIGURE 99

GGCTGACCGTGCTACATTGCCTGGAGGAAGCCTAAGGAACCCAGGCATCCAGCTGCCACGCCCTG
AGTCCAAGATTCTTCCCAGGAACACAAACGTAGGAGACCCACGCTCTGGAAAGCACCAGCCTTA
TCTCTCACCTTCAAGTCCCCTTCTCAAGAATCCTCTGTTCTTGCCTCTAAAGTCTGGTAC
ATCTAGGACCCAGGCATCTGCTTCCAGCCACAAAGAGACAGATGAAGATGCAGAAAGGAAATG
TTCTCCTATGTTGGTCTACTATTGCATTAGAAGCTGCAACAAATTCCAATGAGACTAGCACC
TCTGCCAACACTGGATCCAGTGTGATCTCCAGTGGAGCCAGCACAGCCACCAACTCTGGTCCAG
TGTGACCTCCAGTGGGTCAAGCACAGCCACCATCTCAGGGTCCAGCGTGACCTCCAATGGGTCA
GCATAGTCACCAACTCTGAGTCCATACAACCTCCAGTGGGATCAGCACAGCCACCAACTCTGAG
TTCAGCACAGCGTCCAGTGGGATCAGCATAGCCACCAACTCTGAGTCCAGCACACACCTCCAGTGG
GGCCAGCACAGCCACCAACTCTGAGTCCAGCACACCCTCAGTGGGCCAGCACAGTCACCAACT
CTGGGTCCAGTGTGACCTCCAGTGGAGCCAGCAGTGCACCAACTCTGAGTCCAGCACAGTGTCC
AGTAGGGCCAGCACTGCCACCAACTCTGAGTCTAGCACACTCTCCAGTGGGCCAGCACAGCCAC
CAACTCTGACTCCAGCACAACCTCCAGTGGGCTAGCACAGCCACCAACTCTGAGTCCAGCACAA
CCTCCAGTGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACAGCCACCAACTCTGAGTCCAG
GCCACCAACTCTGAGTCCAGCACAACCTCCAGTGGGCCAGCACAGCCACCAACTCTGAGTCCAG
AACGACCTCCAATGGGCTGGCACAGCCACCAACTCTGAGTCCAGCACGACCTCCAGTGGGCCA
GCACAGCCACCAACTCTGACTCCAGCAGTGTCCAGTGGGCCAGCACAGCCACCAACTCTGAG
TCCAGCACGACCTCCAGTGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACGACCTCCAGTGG
GGCTAGCACAGCCACCAACTCTGACTCCAGCACAACCTCCAGTGGGCCGGCACAGCCACCAACT
CTGAGTCCAGCAGTGTCCAGTGGGATCAGCACAGTCACCAATTCTGAGTCCAGCACACCCCTCC
AGTGGGCCAACACAGCCACCAACTCTGAGTCCAGTACGACCTCCAGTGGGCCAACACAGCCAC
CAACTCTGAGTCCAGCAGTGTCCAGTGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACAA
CCTCCAGTGGGTCAAGCACAGCCACCAACTCTGAGTCCAGCACAACCTCCAGTGGGCTAGCACA
GCCACCAACTCTGACTCCAGCACAACCTCCAGTGGGCCAGCACAGCCACCAACTCTGAGTCTAG
CACAGTGTCCAGTGGGATCAGCACAGTCACCAATTCTGAGTCCAGCACAACCTCCAGTGGGCCA
ACACAGCCACCAACTCTGGTCCAGTGTGACCTCTGCAGGCTCTGGAACAGCAGCTGTGACTGGA
ATGCACACAACCTCCATAGTCATCTACTGCAGTGAGTGAGGCAAAGCCTGGTGGCCCTGGT
GCCGTGGAAATCTTCTCATACCCCTGGTCTCGGTTGTGGCGGCGTGGGCTTTGCTGGC
TCTTCTTCTGTGTGAGAACAGCCTGTCCTGAGAAACACCTTAACACAGCTGTCTACCACCT
CATGGCCTCAACCCTGGCCTGGTCCAGGCCCTGGAGGGATCATGGAGCCCCCACAGGCCAG
GTGGAGTCCTAAGTGGTCTGGAGGAGACAGTATCATCGATAGCCATGGAGATGAGCGGGAGGA
ACAGCGGCCCTGAGCAGCCCCGGAAGCAAGTGCCGCATCTTCAGGAAGGAAGAGACCTGGCA
CCCAAGACCTGGTTCTTCATTCTACCCAGGAGACCCCTCCCAGCTTGTGAGATCCTGAA
AATCTTGAAGAAGGTATTCTCACCTTCTGCTTACCAAGACACTGGAAAGAGAATACTATAT
TGCTCATTAGCTAAGAAATAACATCTCATCTAACACACAGACAAAGAGAAGCTGTGCTG
CCCCGGGGTGGGTATCTAGCTGAGATGAACCTAGTTATAGGAGAAAACCTCCATGCTGGACTC
CATCTGGCATTCAAAATCTCACAGTAAAATCCAAAGACCTCAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAA

FIGURE 100

MKMQKGNVLLMFGLLLHLEAATNSNETSTSANTGSSVISSGASTATNSGSSVTSSGVSTATISGS
SVTSNGVSIVTNSEFHTTSSGISTATNSEFSTASSGISIATNSESSTSSGASTATNSESSTPSS
GASTVTNSGSSVTSSGASTATNSESSTVSSRASTATNSESSTLSSGASTATNSDSSTTSSGASTA
TNSESSSTSSGASTATNSESSTVSSRASTATNSESSTTSSGASTATNSESRTTSNGAGTATNSES
STTSSGASTATNSDSSTVSSGASTATNSESSTTSSGASTATNSESSTTSSGASTATNSDSSTTSS
GAGTATNSESSTVSSGISTVTNSESSTPSSGANTATNSESSTTSSGANTATNSESSTVSSGASTA
TNSESSSTSSGVSTATNSESSTTSSGASTATNSDSSTSSEASTATNSESSTVSSGISTVTNSES
STTSSGANTATNSGSSVTSAGSGTAALTGMHTTSHSASTAVSEAKPGGSLVPWEIFLITLVSVA
AVGLFAGLFFCVRNSLRLNTFNTAVYHPHGLNHGLGPAGGNHGAPHRPRWSPNWFWRPVSSI
AMEMSGRNSGP

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 510-532

FIGURE 101

GGCCGGACGCCTCCCGT TACGGATGAATT AACGGCGGGT CCGCACGGAGGT GTGACCCCTA
CGGAGCCCCAGCTTGC CCACGCACCCC ACTCGGC GTCGCGCGT GCCCTGCTGT CACAGGTG
GGAGGGCTGGA ACTATCAGGCTGAAAAACAGAGTGGGTACTCTCTGGGAAGCTGGCAACAAAT
GGATGATGTGATATATGCATTCCAGGGAAAGGGAAATTGTGGTGCTTCTGAACCCATGGTCAATT
AACGAGGCAGTTCTAGCTACTGCACGTACTTCATAAAGCAGGACTCTAAAAGCTTGGAAATCAT
GGTGT CATGGAAAGGGATT TACTTATACTGACTCTGTTGGGAAGCTTTGGAAAGCATT
TCATGCTGAGTCCCTTTACCTTGTGTTGATGTTGAAACCCATCTGGTATCGCTGGATCAACAAAC
CGCCTTGTGGCAACATGGCTCACCTACCTGTGGCATTATTGGAGACCATGTTGGTGTAAAAGT
GATTATAACTGGGGATGCATTGTTCTGGAGAAAGAAGTGT CATTATCATGAACCATCGGACAA
GAATGGACTGGATGTTCTGTGGAATTGCCTGATGCGATATAGCTACCTCAGATTGGAGAAAATT
TGCCTCAAAGCGAGTCTCAAAGGTGTTCTGGATTGGTGGCCATGCAGGCTGCTGCCTATAT
CTTCATTCA TAGGAAATGGAAGGATGACAAGAGCCATTCGAAGACATGATTGATTACTTTGTG
ATATTCA CGAACCACTTCAACTCCTCATATTCCAGAAGGGACTGATCTCACAGAAAACAGCAAG
TCTCGAAGTAATGCATTGCTGAAAAAAATGGACTTCAGAAATATGAATATGTTTACATCCAAG
AACTACAGGCTTACTTTGTGGTAGACCGTCTAAGAGAAGGTAAAGAACCTTGATGCTGTCCATG
ATATCACTGTGGGTATCCTCACAAACATTCCCTCAATCAGAGAACCTCCCTCCAAGGAGACTT
CCCAGGGAAATCCACTTCACGTCCACCGGTATCCAATAGACACCCTCCCCACATCCAAGGAGGA
CCTTCAACTCTGGTGCACAAACGGTGGAGAGAAAGAAGAGAGGGCTGCCTTCTATCAAG
GGGAGAAGAATT TTTACCGGACAGAGTGT CATTCCACCTGCAAGTCTGAACTCAGGGTC
CTTGTGGTCAAATTGCTCTATACTGTATTGGACCCCTGTTAGCCCTGCAATGTGCCTACTCAT
ATATTGTACAGTCTGTTAAGTGGTATT TATAATCACCATTGTAATCTTGCTGCAAGAGA
GAATATTGGTGGACTGGAGATCATAGAACTTGCATGTTACCGACTTTACACAAACAGCCACAT
TTAAATTCAAAGAAAATGAGTAAGATTATAAGGTTGCCATGTGAAAACCTAGAGCATATTTG
GAAATGTTCTAACCTTCTAAGCTCAGATGCATTGCACTGACTATGTCGAATATTCTTACT
GCCATCATTATTGTTAAAGATATTGCACTTAATTGTGGAAAAATATTGCTACAATT
TTAATCTCTGAATGTAATTGATACTGTGTACATAGCAGGGAGTGATCGGGGTGAAATAACTT
GGGCCAGAATATTAAACAATCATCAGGCTTTAAA

FIGURE 102

MHSRGREIVVLLNPWSINEAVSSYCTYFIKQDSKSGIMVSWKGIFYFILTLFWGSFFGSIFMLSP
FLPLMFVNPSWYRWINNRLVATWLTPVALLETMFGVKVIITGDAFVPGERSVIIMNHRTRMDWM
FLWNCLMRYSYLRLEKICLKASLGPGFGWAMQAAAYIFIHRKWKDDKSHFEDMIDYFCDIHEP
LQLLIFPEGTDLTENSRSNAFAEKNGLQKYEYVLHPRTGFTVVDRLREGKNLDAVHDITVA
YPHNIPQSEKHLQLQGDFPREIHFHVRYPIDTLPTSKEDLQLWCHKRWEEKERLRSFYQGEKNF
YFTGQSVIPPCKSELRLVVKLLSILYWTLFSPAMCLLIYLYSLVKWYFIITIVIFVLQERIFGG
LEIIELACYRLLHKQPHLNSKKNE

Important features of the protein:

Signal peptide:

amino acids 1-22

Transmembrane domains:

amino acids 44-63, 90-108, 354-377

FIGURE 103

CGGCTCGAGCGGCTCGAGTGAAGAGCCTCTCCACGGCTCCTGCGCCTGAGACAGCTGGCCTGACC
TCCAAATCATCCATCCACCCCTGCTGTATCTGTTTCATAGTGTGAGATCAACCCACAGGAATA
TCCATGGCTTTGTGCTCATTGGTTCTCAGTTCTACGAGCTGGTGTAGGACAGTGGCAAGT
CACTGGACCGGGCAAGTTGTCCAGGCCTGGTGGGGGAGGACGCCGTGTTCTGCTCCCTCT
TTCCTGAGACCAGTGCAGAGGCTATGGAAGTGCCTTCAGGAATCAGTCCATGCTGTGGTC
CACCTCTACAGAGATGGGAAGACTGGGAATCTAACGAGATGCCACAGTATCGAGGGAGAACTGA
GTTTGTGAAGGACTCCATTGCAGGGGGCGTGTCTCTAACGGCTAAAAAACATCACTCCCTCGG
ACATCGGCCTGTATGGGTGCTGGTCAGTTCCAGATTACGATGAGGAGGCCACCTGGGAGCTG
CGGGTGGCAGCACTGGCTCACTCCTCTCATTCCATCGTGGATATGTTGACGGAGGTATCCA
GTTACTCTGCCTGTCCCTCAGGCTGGTCCCCCAGCCCACAGCCAAGTGGAAAGGTCCACAAGGAC
AGGATTGTCTTCAGACTCCAGAGCAAATGCAGATGGGTACAGCCTGTATGATGTGGAGATCTCC
ATTATAGTCCAGGAAAATGCTGGGAGCATATTGTGTTCCATCCACCTTGTGAGCAGAGTCATGA
GGTGGAAATCCAAGGTATTGATAGGAGAGACGTTTTCCAGCCCTCACCTGGCGCTGGCTTCTA
TTTACTCGGGTTACTCTGTGGTGCCTGTGTGGTGTGTCATGGGATGATAATTGTTCTTC
AAATCCAAAGGGAAAATCCAGGCAGGAACTGGACTGGAGAAGAAAGCACGGACAGGAGAAATTGAG
AGACGCCCGGAAACACGCAGTGGAGGTGACTCTGGATCCAGAGACGGCTCACCGAAGCTCTGCG
TTCTGATCTGAAAATGTAACCCATAGAAAAGCTCCCCAGGAGGTGCCTCACTCTGAGAAGAGA
TTTACAAGGAAGAGTGTGGTGGCTTCAGGGTTCCAAGCAGGGAGACATTACTGGGAGGTGGA
CGTGGGACAAAATGTAGGGTGGTATGTGGAGTGTGTCGGATGACGTAGACAGGGGAGAAACA
ATGTGACTTGTCTCCAACAATGGGTATTGGGTCTCAGACTGACAACAGAACATTGTATTTC
ACATTCAATCCCCATTATCAGCCTCCCCCAGCACCCCTCCTACACGAGTAGGGTCTTCC
GGACTATGAGGGTGGGACCATCTCCTCTCAATACAAATGACCAGTCCCTTATTATACCCTGC
TGACATGTCAGTTGAAGGCTTGTGAGACCCATATCCAGCATGCGATGTATGACGAGGAAAAG
GGGACTCCCATATTCATATGTCAGTGTCTGGGGATTGAGACAGAGAACCCCTGCTAAAGGGC
CCCACACCACAGACCCAGACACAGCCAAGGGAGAGTGTCCCCGACAGGTGGCCCCAGCTTCC
CCGGAGCCTGCGCACAGAGAGTCACGCCCAACTCTCCTTAGGGAGCTGAGGTTCTCTGCC
TGAGCCCTGCAGCAGCGCAGTCACAGCTCCAGATGAGGGGGATTGGCCTGACCCGTGGGAG
TCAGAAGCCATGGCTGCCCTGAAGTGGGAGGAATAGACTCACATTAGTTAGTTGTGAAA
CTCCATCCAGCTAACGATCTGAACAAGTCACAACCTCCAGGCTCCTCATTGCTAGTCACGG
ACAGTGATTCTGCCTCACAGGTGAAGATTAAAGAGACAACGAATGTGAATCATGCTTGAGGTT
TGAGGGCACAGTGTGCTAATGATGTGTTTATATTACATTCCCACCAAACTCTGTT
TGCTTATTCCACATTAATTACTTTCTATACCAAATACCCATGGAATAGTTATTGAACACC
TGCTTGTGAGGCTCAAAGAATAAAGAGGGAGGTAGGATTTCACTGATTCTATAAGCCCAGCAT
TACCTGATACCAAAACAGGCAAAGAAAACAGAAGAAGAGGAAGGAAAACACTACAGGTCCATATCC
CTCATTAACACAGACACAAAAATTCTAAATAAAATTAAACAAATTAAACTAAACAAATATTAA
AAGATGATATATAACTACTCAGTGTGGTTGTCCCACAAATGCAGAGTTGGTTAATATTAAAT
ATCAACCAGTGTAAATTCAAGCTAAAGTAAAAAGAAAACCATAAAAAAA

FIGURE 104

MAFVLILVLSFYELVSGQWQVTGPGKFVQALVGEDAVFSCSLFPETSAEAMEVRFFRNQFHAVVH
LYRAGEDWESKQMPQYRGRTEFKDSIAGGRVSLRLKNITPSDIGLYGCWFSSQIYDEEATWELR
VAALGSLPLISIVGYVDGGIQLLCLSSGWFPQPTAKWKGPQGQDLSSDSRANADGYSLYDVEISI
IVQENAGSILCSIHLAEQSHEVESKVLIGETFFQPSPWRLASILLGLLCGALCGVVMGMIIVFFK
SKGKIQAELDWRRKHGQAEELRDARKHAVEVTLDPETAHPKLCVSDLKTVTHRKAPQEVPHSEKRF
TRKSVVASQGFQAGRHYWEVDVGQNVGWWVGVCRDDVRGKNNVTLSNNNGYWVLRLTTEHLYFT
FNPHFISLPPSTPPTRGVFLDYEGGTISFFNTNDQSLIYTLLTCQFEGLLRPIQHAMYDEEKG
TPIFICPVSWG

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 131-150, 235-259

FIGURE 105

CCTTCACAGGACTCTCATTGCTGGTGGCAATGATGTATCGGCCAGATGTGGTAGGGCTAGGAAAAGAG
TTTGGTGGAACCCCTGGTTATCGGCCTCGTCATCTCATATCCCTGATTGTCTGGCAGTGTGCATTGGA
CTCACTGTTCAATTATGTGAGATATAATCAAAAGAACCTACAATTACTATAGCACATTGTCAATTACAAC
TGACAAACTATATGCTGAGTTGGCAGAGAGGCTCTAACAAATTTACAGAAATGAGCCAGAGACTTGAAT
CAATGGTAAAAATGCATTTATAAATCTCCATTAAGGGAGAATTGTCAAGTCTCAGGTTATCAAGTTC
AGTCAACAGAACATGGAGTGTGGCTCATATGCTGTTGATTGTAGATTCACTCTACTGAGGATCCTGA
AACTGTAGATAAAATTGTTCAACTTGTAACTGAAAGCTGCAAGATGCTGTAGGACCCCTAAAGTAG
ATCCTCACTCAGTTAAATTAAAAATCAACAAGAACAGAACAGCTATCTAAACCATTGCTGCAG
ACACGAAGAAGTAAACTCTAGGTCAAGACTCAGGATCGTGGTGGACAGAAGTAGAAGAGGGTGAATG
GCCCTGGCAGGCTAGCCTGCAGTGGATGGAGTCATCGCTGTGGAGCAACCTTAATTGCCACATGGC
TTGTGAGTGCTGCTCACTGTTAACATATAAGAACCCGCCAGATGGACTGCTCCTTGGAGTAACA
ATAAAACCTCGAAATGAAACGGGTCTCCGGAGAATAATTGTCCATGAAAAATACAAACACCCATCACA
TGACTATGATATTCTCTTGAGAGCTTCTAGCCCTGTTCCACACAAATGCAGTACATAGAGTTGTC
TCCCTGATGCATCCTATGAGTTCAACCAGGTGATGTGATGTTGTGACAGGATTGGAGCACTGAAAAT
GATGGTTACAGTCAAATCATCTCGACAAGCACAGGTGACTCTCATAGACGCTACAACCTGCAATGAACC
TCAAGCTTACAATGACGCCATAACTCCTAGAATGTTATGTGCTGGCTCCTAGAAGGAAAACAGATGCAT
GCCAGGGTGAECTGGAGGACCCTGGTTAGTCAGATGCTAGAGATATCTGGTACCTGCTGGAATAGTG
AGCTGGGAGATGAATGTGCAAAACCAACAAGCCTGGTTATACTAGAGTTACGGCCTGCGGGACTG
GATTACTTCAAAACTGGTATCTAAGAGACAAAGCCTCATGGAACAGATAACATTGTTGTTGTTG
GGTGTGGAGGCCATTAGAGATACAGAATTGGAGAAGACTTGCAAAACAGCTAGATTGACTGATCTCA
ATAAAACTGTTGCTTGATGCATGTATTCTTCCAGCTCTGTCGGCACGTAAGCATCCTGCTTCTGCCA
GATCAACTCTGTCATCTGTGAGCAATAGTTGAAACTTATGTACATAGAGAAATAGATAATACAATATTAC
ATTACAGCCTGTATTCTAGAAGTTGTCAGAATTGACTGACATAAATTGTAAT
GCATATATACAATTGAAGCACTCCTTCTCAGTCCTCAGCTCCTCATTCAAGCAAATATCCATT
TCAAGGTGCAGAACAGGAGTGAAAGAAAATATAAGAAGAAAAATCCCTACATTGACAGAA
AAGTATTAGGTGTTCTAGTGAATATTAGAAATGATCATATTCAATTGAAAGGTCAAGCAAAGACA
GCAGAACATCAACTCATCTAGAAGTATGGAAACTAAGTTAAGGAAGTCCAGAAAGAACCAAG
ATATATCCTTATTTCATTCAAACAACTACTATGATAATGTGAAGAAGATTCTGTTGACCT
ATAATAATTACAAACTCATGCAATGTACTTGTCTAAGCAAATTAAAGCAAATTATTAACATTG
TTACTGAGGATGTCAACATATAACAATAAAATATAACACCCCA

FIGURE 106

MMYRPDVVRARKRVCWEPWVIGLVIFISLIVLAVCIGLTVHYVRYNQKKTYNYYSTLSFTTDKLY
AEFGREASNNFTEMSRQLESVKNAFYKSPLREFVKSQVIKFQQKHGVLAHMLLICRFHSTED
PETVDKIVQLVLHEKLQDAVGPPKVDPHSVKIKKINKTETDSYLNHCCGTRRSKTLQSLRIVGG
TEVEEGEWPWQASLQWDGSHRCGATLINATWLVSAAHCFETYKNPARWTASFGVTIKPSKMKGRL
RRIIVHEKYKHPSHDYDISLAEELSSPVPTNAVHRVCLPDASYEFQPGDVMFVTGFGALKNDGYS
QNHLRQAQVTLIDATTNEPQAYNDAITPRMLCAGSLEGKTDACQGDGGPLVSSDARDIWYLAG
IVSWGDECAPNKPGVYTRVTALRDWITSKTGI

Transmembrane domain:
amino acids 21-40 (type II)

FIGURE 107

AGAGAAAGAACGTCTCCAGCTGAAGCCAATGCAGCCCTCCGGCTCTCCGCGAAGAACGTTCCCTG
CCCCGATGAGCCCCCGCCGTGCGTCCCCACTATCCCCAGGCAGGGCTGGGGCACCAGGCCCAGC
GCCGACGATCGCTGCCGTTGCCCTGGGAGTAGGATGTTGAAAGGATGGGGCTTCTCCCT
ACGGGGCTCACAATGCCAGAGAACGATTCCGTGAAGTGTCTGCGCTGCCTGCTCTACGCCCTCAA
TCTGCTCTTGTTAATGTCCATCAGTGTGGCAGTTCTGCTGGATGAGGGACTACCTAA
ATAATGTTCTCACTTAACAGCAGAACGAGGGTAGAGGAAGCAGTCATTGACTTACTTCCT
GTGGTTCATCCGGTCATGATTGCTGTTGCTGTTCTTATCATTGTTGGGGATGTTAGGATATTG
TGGAACGGTGAAAAGAAATCTGTTGCTCTTGACATGGTACTTGGAAAGTTGCTGTCATTTCT
GTGTTAGAACTGGCTGTGGCTTGGACATATGAACAGGAACATTGGTTCCAGTACAATGGTCA
GATATGGTCACTTGAAAGCCAGGATGACAAATTATGGATTACCTAGATATCGGTGGCTTACTCA
TGCTTGAATTTCAGAGAGAGTTAAGTGTGTGGAGTAGTATATTCACTGACTGGTTGG
AAATGACAGAGATGGACTGGCCCCAGATTCTGCTGTGTTAGAGAAATTCCAGGATGTTCCAAA
CAGGCCACCAGGAAGATCTCAGTGACCTTATCAAGAGGGTTGGGAAGAAAATGTATTCTT
TTTGAGAGGAACCAAACAACTGCAGGTGCTGAGGTTCTGGGAATCTCCATTGGGGTACACAAA
TCCTGGCCATGATTCTCACCATTACTCTGCTCTGGCTCTGTATTATGATAGAACGGGAGCCTGGG
ACAGACCAAATGATGTCCTGAAGAATGACAACACTCTCAGCACCTGTGATGTCCTCAGTAGAACT
GTTGAAACCAAGCCTGTCAAGAATCTTGAACACACATCCATGGCAAACAGCTTAATACACACT
TTGAGATGGAGGAGTTAAAAAGAAATGTCACAGAAGAAACCACAAACTGTTTATTGGACT
TGTGAATTGAGTACATACTATGTGTTGAGAAATATGTTAGAAATAAAATGTTGCCATAAAA
TAACACCTAACATATACTATTCTATGCTTAAAGGATGGAAAAGTTCATGTCATAAGTC
ACCACCTGGACAATAATTGATGCCCTAAATGCTGAAGACAGATGTGATACCCACTGTGAGCC
TGTGTATGACTTTACTGAACACAGTTATGTTGAGGCAGCATGGTTGATTGATGCTTCCGCA
TCCATGCAAACGAGTCACATATGGTGGACTGGAGCCATAGTAAAGGTTGATTACTTCTACCAA
CTAGTATATAAAAGTACTAATTAAATGCTAACATAGGAAGTTAGAAAATACTAATAACTTTATT
CTCAGCGATCTTCTGATGCTAACATAATTATGCTAACATATTGAGAAAACCTTCAATATTGGTACT
ACCTAAATGTGATTGCTGGTTACTAAATATTCTTACCACTTAAAGAGCAAGCTAACACAT
TGTCTTAAGCTGATCAGGGATTGTTGATATAAGTCTGTTAAATCTGTATAATTCACTGCGAT
TTCAGTTCTGATAATGTTAAGAATAACCATTATGAAAGAAAATTGTCCTGTATAGCATCATT
ATTTTAGCCTTCTGTTAATAAGCTTACTATTCTGCTCTGGCTTATATTACACATATAAC
TGTTATTTAAACTTAACCACTAATTGAAATTACCACTGATAGCATAGGAATCATTATTC
AGAATGTAGTCTGGCTTGTAGAAGTATTAAGAAAATTGACATAACTTAGTTGATTGAGA
AAGGACTTGTATGCTTTCTCCAAATGAAGACTCTTTGACACTAAACACTTTAAAAA
GCTTATCTTGCCTCTCCAAACAGAACAGCAATAGTCTCCAAGTCAATATAAATTCTACAGAAA
TAGTGTCTTTCTCCAGAAAATGCTGTGAGAATCATTAAACATGTGACAATTAGAGATT
CTTGTTTATTTCACTGATTAATATACTGTGGCAAATTACACAGATTATTAATTTTACAA
GAGTATAGTATATTGAAATGGAAAAGTGCATTTACTGTATTTGTGTTAGAGATT
TTCTCAGAATATGAAAGAAAATTAAATGTGTCATAAAATTTCTAGAGAGTAA

FIGURE 108

MAREDSVKCLRCLLYALNLLFWLMSISVLAWSAWMRDYLNNVLTLTAETRVEEAVILTYFPVVHP
VMIAVCCFLIIVGMLGYCGTVKRNLNLLAWYFGSLLVIFCVELACGVWTYEQELMVPVQWSDMVT
LKARMTNYGLPRYRWLTHAWNFFQREFKCCGVVYFTDWLEMTEMDWPPDSCCVREFPGCSKQAHQ
EDLSDLYQEGCGKKMYSFLRGTKQLQVLRLGIISIGVTQILAMILTTITLLWALYYDRREPGTDQM
MSLKNDNSQHLSCPSVELLKPSLSRIFEHTSMANSFNTHFEMEEL

Signal peptide:

amino acids 1-33

Transmembrane domains:

amino acids 12-35, 57-86, 94-114, 226-248

FIGURE 109

CCAAGGCCAGAGCTGTGGACACCTTATCCCACTCATCCTCATCCTCTGATAAAGCCCCTACCACTGCT
GATAAAAGTCTTCTCGTGAGAGCCTAGAGGCCTTAAAAAAAAGTGTGAAAGAGAAGGGACAAAGGAACA
CCAGTATTAAGAGGATTTCCAGTGTTCAGTGGCAGTGGTCCAGAAGGATGCCTCCATTCTGCTCTCACCTG
CCTCTTCATCACAGGCACCTCCGTGTCACCCGTGCCCTAGATCCTGTTCTGCTTACATCAGCCTGAATGAGC
CCTGGAGGAACACTGACCACCAGTGGATGAGTCTCAAGGTCTCCTCTATGTGACAACCAGTGAATGGGAG
TGGTACCACTTCACGGGCATGGCGGGAGATGCCATGCCTACCTTCTGCATACCAGAAAACCACTGTGGAACCCA
CGCACCTGTCTGGCTCAATGGCAGCCACCCCTAGAAGGCACGGCATTGTGCAACGCCAGGCTGTGCCAGCT
TCAATGGGAACTGCTGTCTGGAACACCACGGTGAAGTCAAGGCTTGCCTGGAGGCTACTATGTGTATCGT
CTGACCAAGCCCAGCGTCTGCTTCCACGTACTGTGGTCAAGGCTTGCCTGGAGGCTACTATGTGTATCGT
CAGCTGCTCAGATAACCAGCGAGTGACATGCCCTCAGGAACACTGTGCTAGGCCCTGACAGGCAGACATGCTTG
ATGAAAATGAATGTGAGCAAAACACGGTGGCTGCAGTGAGATCTGTGAAACCTCAAAACTCCTACCGCTGT
GAGTGTGGGTTGGCGTGTGCTAAGAAGTGTGAAAGACTTGTGAAAGACGTTGAAGGATGCCACAATAACAA
TGGTGGCTGCAGCCACTCTGCCTTGGATCTGAGAAAGGCTACCAAGTGTGAATGTCCCCGGGCCTGGTGCTGT
CTGAGGATAACCACACTTGCAAGTCCCTGTGTCAGGAAATGCCATTGAAGTGAACATCCCCAGGGAG
CTGGTTGGTGGCCTGGAGCTCTCCTGACCAACACCTCTGCCGAGGAGTGTCCAACGGCACCCATGTCAACAT
CCTCTCTCTCAAGACATGTGGTACAGTGGTCATGTGGTGAATGACAAGATTGTGCCAGCAACCTCGTGA
CAGGTCTACCCAAGCAGACCCCCGGGAGCAGCGGGACTTCATCATCCGAACCAGCAAGCTGCTGATCCCAGG
ACCTGCGAGTTCCACGCCGTACACCATTCTGAAGGATACTTCCAAACCTCGAAACTCCCCACTGGAAAT
CATGAGCCGAAATCATGGATCTTCCATTCACTCTGGAGATCTCAAGGACAATGAGTTGAAGAGCCTTACC
GGGAAGCTCTGCCACCCCAAGCTCGTACTCCCTCTACTTGGCATTGAGCCGTGGTCACGTGAGCGGC
TTGGAAAGCTTGGTGGAGAGCTGTTGCCACCCCACTCCAAGATGACCGAGGTCTGAAATACTACCTCAT
CCGGGATGGCTGTGTTCAAGATGACTCGGTAAAGCAGTACACATCCCGGATCACCTAGCAAAGCACTTCCAGG
TCCCTGCTTCAAGTTGTGGCAAAGACCAAGGAAGTGTGTTCTGACTGCCGGTTCTTGTCTGTGGAGTG
TTGGACGAGCGTCCCGCTGTGCCACGGGTTGCCACCGCGAATGCGTGTGGGGCAGGAGGAGAGGACTCAGC
CGGTCTACAGGGCCAGACGCTAACAGCGGCCGATCCGATCGACTGGGAGGACTTAGTCGTAGCCATACCTC
GAGTCCCTGCATTGGACGGCTCTGCTCTTGAGCTTCTCCCCCACC GCCCTTAAGAACATCTGCCAACAGC
TGGGTCAGACTCACACTGTGAGTTCAAGACTCCCAGCACCAACTCACTCTGATTCTGGTCCATTCACTGGGCA
CAGGTACAGCACTGCTGAACAATGTGGCTGGTGGGTTCATCTTCTAGGGTTGAAAACAAACTGTCCA
CCCAGAAAGACACTCACCCATTCCCTCATTTCTTCAACTAAACCTCGTGTATGGTCAATCAGAC
CACAAAATCAGAAGCTGGGTATAATATTCAAGTTACAAACCCCTAGAAAATTAAACAGTTACTGAAATTATGA
CTTAAATACCAATGACTCCTAAATATGAAATTATAGTTACCTTGAAATTCAATTCAAATGCAGACTAA
TTATAGGAATTGGAAGTGTATCAATAAAACAGTATATAATT

FIGURE 110

MPPFLLTCLFITGTSVSPVALDPCSAYISLNEPWRNTDHQLDESQGPPLCDNHVNGEWYHFTGMAGDAMP
TFCIPEHCGTHAPVWLNGSHPLEGDGIVQRQACASFNGNCCLWNTTVEVKACPGGYYVYRLTKPSVCFHV
YCGHFYDICDEDCHGSCSDTSECTCAPGTVLGPDRQTCFDENECEQNNGCSEICVNLKNSYRCECGVGRV
LRSDGKTCEDVEGCHNNNGCSEKGYQCECPRLVLSEDNHTCQVPVLCKSNAIEVNIPRELVGG
LELFILTNTSCRGVSNGTHVNILFSLKTCGTVVVVNDKIVASNLVTGLPKQTPGSSGDFIIRTSKLLIPVT
CEFPRLYTISEGYVPNLRNSPLEIMSRNHGIFPFTLEIFKDNEFEEPYREALPTLKLRSLSYFGIEPVVHV
SGLESLVESCFATPTSKIDEVLKYLYIRDGCVSDDSVKQYTSRDHLAKHFQVPVFKFVGKDHKEVFLHCRV
LVCGVLDERSRCAQGCHRRMRRGAGGEDSAGLQGQTLTGGPIRIDWED

Important features of the protein:

Signal peptide:

amino acids 1-16

N-glycosylation sites.

amino acids 89-93, 116-120, 259-263, 291-295, 299-303

Tyrosine kinase phosphorylation sites.

amino acids 411-418, 443-451

N-myristoylation sites.

amino acids 226-232, 233-239, 240-246, 252-258, 296-302, 300-306,
522-528, 531-537

Aspartic acid and asparagine hydroxylation site.

amino acids 197-209

ZP domain proteins.

amino acids 431-457

Calcium-binding EGF-like proteins.

amino acids 191-212, 232-253

FIGURE 111

GAGAGAGGCAGCAGCTGCTCAGCGGACAAGGATGCTGGCGTGAGGGACCAAGGCCTGCCCTGCACTCGG
GCCTCCTCCAGCCAGTGCTGACCAGGGACTCTGACCTGCTGCCAGCCAGGACCTGTGTGGGGAGGCCT
CCTGCTGCCTTGGGTGACAATCTCAGCTCCAGGCTACAGGGAGACCGGGAGGATCACAGAGCCAGCATGT
TACAGGATCCTGACAGTGATCAACCTCTGAACAGCCTCGATGTCAAACCCCTGCGCAAACCCGTATCCCC
ATGGAGACCTTCAGAAAGGTGGGATCCCCATCATCATAGCACTACTGAGCCTGGCGAGTATCATCATTGT
GGTTGTCCATCAAGGTGATTCTGGATAAAACTACTTCCTCTGCCAGCCTCTCCACTTCATCCGA
GGAAGCAGCTGTGTGACGGAGAGCTGGACTGTCCCTGGGGAGGACGAGGAGCACTGTGTCAAGAGCTTC
CCCGAAGGGCTGCAGTGGCAGTCCGCCTCTCCAAGGACCGATCCACACTGCAGGTGCTGGACTCGGCCAC
AGGGAACTGGTCTGCCTGTTGACAACCTCACAGAAGCTCGCTGAGACAGCCTGTAGGCAGATGG
GCTACAGCAGAGCTGTGGAGATTGGCCAGACAGCAGGATCTGGATGTTGAAATCACAGAAAAGCCAG
GAGCTTCGCATGCGGAACACTCAAGTGGCCCTGTCTCAGGCTCCCTGGCTCCCTGCACTGTCTGCCCTG
TGGGAAGAGCCTGAAGACCCCCCGTGTGGTGGTGGGGAGGAGGACCTCTGTGGATTCTTGGCCTTGGCAGG
TCAGCATCCAGTACGACAAACAGCACGTCTGTGGAGGGAGCATTGGACCCCCACTGGTCCTCACGGCA
GCCCACTGCTTCAGGAACATACCGATGTGTTCAACTGGAAGGTGCGGGCAGGCTCAGACAAACTGGCAG
CTTCCCATCCCTGGCTGTGGCAAGATCATCATCATTGAATTCAACCCATGTACCCAAAGACAATGACA
TCGCCCTCATGAAGCTGCAGTCCCACACTTCTCAGGCACAGTCAGGCCATCTGTCTGCCCTTCTT
GATGAGGAGCTCACTCCAGCCACCCACTCTGGATCATTGGATGGGCTTACGAAGCAGAATGGAGGGAA
GATGCTGACATACTGCTGCAGCGTCAGTCCAGGTATTGACAGCACACGGTGCAATGCAGACGATGCGT
ACCAGGGGAAGTCACCGAGAAGATGATGTGCAGGCATCCCGAAGGGGTGTGGACACCTGCCAGGGT
GACAGTGGTGGCCCTGATGTACCAATCTGACCAAGGTGGCATGTGGTGGCATCGTAGCTGGCTATGG
CTGGGGGGCCCGAGCACCCAGGGAGTACACCAAGGTCTCAGCCTATCTCAACTGGATCTACAATGTCT
GGAAGGCTGAGCTGTAATGCTGCCCTTGCACTGGCTGGAGGCCCTCCCTGCCCTGCCACCT
GGGATCCCCAAAGTCAGACACAGAGCAAGAGTCCCTGGGTACACCCCTCTGCCACAGCCTCAGCAT
TTCTGGAGCAGCAAAGGGCTCAATTCTGTAAGAGACCCCTCGCAGCCCAGAGGCCAGAGGAAGTCA
GCAGCCCTAGCTGCCACACTTGGTCTCCAGCATCCCAGGGAGAGACACAGCCCACGTAAAGGTCT
CAGGGTATTGCTAAGCCAAGAAGGAACCTTCCCACACTACTGAATGGAAGCAGGCTGTCTTGTAAAAGCC
CAGATCACTGTGGCTGGAGAGGAGAAGGAAAGGGTCTGCCAGCCCTGTCCGTCTCACCCATCCCCAA
GCCTACTAGAGCAAGAACCAAGTTGTAATATAAAATGCACTGCCCTACTGTTGGTATGACTACCGTTACCT
ACTGTTGTCATTGTTATTACAGCTATGCCACTATTATTAAAGAGCTGTGTAACATCTGGCAAAAAAAA
AAAA

FIGURE 112

MLQDPDSDQPLNSLDVKPLRKPRIPMETFRKVGIP III ALLSLASIIIVVLIKVILDKYYFLCG
QPLHFIPRKQLCDGELDCPLGEDEEHCVKSFPEGPAVAVRLSKDRSTLQVLDSATGNWFSACFDN
FTEALAAETACRQMGYSRAVEIGPDQDLDVVEITENSQELMRMRNSSGPCLSGSLVSLHCLACGKSL
KTPR VVGEEASVDSWPWQVSIQYDKQHVC GGSILD PHWVL TAAHCFRKHTDVF NWKVRAGSDKL
GSF PSLAVAK IIIIEFNPMYPKDNDIALMKLQFPLTFSGTVRPICL PFFDEELTPATPLWIIGWG
FTKQNGGKMSDILLQASVQVIDSTRCNADDAYQGEVTEKMMCAGIPEGGVDT CQGD SGGPLMYQS
DQWHVVGIVSWG YGCGGPSTPGVYTKV SAYLNWIYNVWKAEL

Transmembrane domain:

amino acids 32-53 (typeII)

FIGURE 113

GGCTGGACTGGAACCTCTGGTCCAAGTGATCCACCCGCCTCAGCCTCCAAAGGTGCTGTGATTA
TAGGTGTAAGCCACCGTGTCTGGCCTCTGAACAACCTTTCACTAAAGCCACAGGAGT
TGAACGTCTAGGATTCTGACTATGCTGTGGCTAGTGCCTACTCCTACCTACATTAAAATC
TGTTTTGTTCTCTGTAACCTAGCCTTACCTCCTAACACAGAGGATCTGTCAGTGTGGCTCT
GCCCAAAACCTGACCTCACTCTGGAACGAGAACAGAGGTTCTACCCACACCGTCCCCTGAAG
CCGGGGACAGCCTCACCTGCTGGCCTCGCTGGAGCAGTGCCTCACCAACTGTCTACGTCT
GGAGGCACTGACTCGGGCAGTGCAGGTAGCTGAGCCTCTGGTAGCTGCCTTCAAGGTGGC
CTTGCCTGGCGTAGAAGGGATTGACAAGCCGAAGATTCAAGGCGATGGCTCCACTGCC
AGGCATCAGCCTTGCTGTAGTCATCAACTGCCCTGGGCCAGGACGGGCCGGACACCTGCTCA
GAAGCAGTGGGTGAGACATCACGCTGCCGCCATCTAACCTTTCATGTCCCTGCACATCACCTG
ATCCATGGGCTAATCTGAACTCTGTCCAAGGAACCCAGAGCTTGAGTGAGCTGTGGCTCAGACC
CAGAAGGGGTCTGCTTAGACCACCTGGTTATGTGACAGGACTTGCAATTCTCCTGAAACATGAGG
GAACGCCGGAGGAAAGCAAAGTGGCAGGGAAAGGAACCTGTGCCAATTATGGTCAGAAAAGATG
GAGGTGTTGGTTATCACAAGGCATCGAGTCTCCTGCATTCACTGGACATGTGGGAAAGGGCTG
CCGATGGCGCATGACACACTCGGGACTCACCTCTGGGCCATCAGACAGCCGTTCCGCCCGAT
CCACGTACCAGCTGCTGAAGGGCAACTGCAGGCCATGCTCTCATCAGCCAGGCAGGCCAAA
TCTGCGATCACCAGCCAGGGCAGCCGCTGGAAAGGAGCAAGCAAAGTGACCAATTCTCCTCCC
CTCCTCCCTCTGAGAGGCCCTCTATGTCCCTACTAAAGCCACCAGCAAGACATAGCTGACAGG
GGCTAATGGCTCAGTGTGGCCAGGAGGTCAAGCACAGGCTGCCAAATCCCCAGGCAAAGGACTGTGTGG
CTCAATTAAATCATGTTCTAGTAATTGGAGCTGTCCCCAAGACCAAAAGGAGCTAGAGCTTGGTT
CAAATGATCTCAAGGGCCCTTATACCCAGGAGACTTGTGATTGAAATTGAAACCCAAATCCA
AACCTAAGAACCAAGGTGCATTAAGAACATCAGTTATTGCCGGGTGTGGCTGTAATGCCAACAT
TTGGGAGGCCAGGGGGTAGATCACCTGAGGTCAAGGAGCTAACAGGCCATGGCAACATGG
TGAAACCCCTGTCTACTAAAAAATACAAAAAAACTAGCCAGGCATGGTGGTGTGCCTGTATC
CCAGCTACTCGGGAGGCTGAGACAGGAGAATTACTTGAACCTGGAGGTGAAGGAGGCTGAGACA
GGAGAATCACTCAGCCTGAGCAACACAGCGAGACTCTGTCAGAAAAAATAAAAAAGAATTA
TGGTTATTGTAA

FIGURE 114

MLWWLVLLLLPTLKSVFCSLVTSLYLPNTEDLSLWLWPKPDLHSGTRTEVSTHTVPSKPGTASPC
WPLAGAVPSPTVSRLEALTRAVQVAEPLGSCGFQGGPCPGRRD

Signal peptide:

amino acids 1-15

FIGURE 115

CAGCAGTGGTCTCTCAGTCCTCTAAAGCAAGGAAAGAGTACTGTGTGCTGAGAGACCATGGCAA
AGAACCTCCAGAGAATTGTGAAGACTGTCACATTCTAAATGCAGAAGCTTTAAATCCAAGAAA
ATATGTAAATCACTTAAGATTGTGGACTGGTGTGTTGCTGCCCTAACTCTAATTGTCT
GTTTGGGGAGCAAGCACTCTGCCGGAGGTACCCAAAAAGCCTATGACATGGAGCACACTT
TCTACAGCAATGGAGAGAAGAAGAAGATTACATGGAAATTGATCCTGTGACCAGAACTGAAATA
TTCAGAACGGAAATGGCACTGATGAAACATTGAAAGTGCACGACTTAAAAACGGATAACTGG
CATCTACTTCGTGGTCTTCAAAAATGTTTATCAAAACTCAGATTAAAGTGAATTCTGAATT
CTGAACCAGAAGAGGAAATAGATGAGAATGAAGAAATTACCAACTTCTTGAACAGTCAGTG
ATTTGGTCCCAGCAGAAAAGCCTATTGAAAACCGAGATTCTTAAAAATTCCAAAATTCTGGA
GATTGTGATAACGTGACCATGTATTGGATCAATCCCCTCTAATATCAGTTCTGAGTTACAAG
ACTTGAGGAGGGAGGAGAAGATCTTCACTTCCGCCAACGAAAAAAAAGGGATTGAACAAAAT
GAACAGTGGTGGTCCCTCAAGTGAAGTAGAGAAGACCCGTACGCCAGACAAGCAAGTGAGGA
AGAACTCCAATAATGACTATACTGAAAATGGAATAGAATTGATCCCCTGCTGGATGAGAGAG
GTTATTGTTGTTACTGCCGTGAGGCAACCGCTATTGCCGCCGTCTGTGAACCTTACTA
GGCTACTACCCATATCCATACTGCTACCAAGGAGGACGAGTCATCTGCGTGTCTGCTTG
TAACTGGTGGTGGCCCGATGCTGGGAGGGCTAATAGGAGGTTGAGCTAAATGCTTAAAC
TGCTGGCAACATATAATAATGCATGCTATTCAATGAATTCTGCCTATGAGGCATCTGCCCT
GGTAGCCAGCTCTCCAGAATTACTGTAGGTAATTCTCTTCATGTTCTAATAAAACTCTACA
TTATCACCAAAAAAAAAAAAAAA

FIGURE 116

MAKNPPENCEDCHILNAEAFSKICKSLKICGLVFGILALTLIVLFWGSKHFPEVPKKAYDME
HTFYSNGEKKKIYMEIDPVTRTEIFRSGNGTDETLEVHDFKNGYTGIYFVGLQKCFIKTQIKVIP
EFSEPEEEIDENEEITTFEQSVIWPVPAEKPIENRDFLKNSKILEICDNVTMYWINPTLISVSE
LQDFEEEGEDLHFPAKEKKGIEQNEQWVVPQVKVEKTRHARQASEEELPINDYTENGIEFDPMED
ERGYCCIYCRRGNRYCRRVCEPLLGYYPYCYQGGRVICRVIMPCNWWVARMLGRV

Important features of the protein:

Signal peptide:

amino acids 1-40

Transmembrane domain:

amino acids 25-47 (type II)

N-glycosylation sites.

amino acids 94-97, 180-183

Glycosaminoglycan attachment sites.

amino acids 92-95, 70-73, 85-88, 133-136, 148-151, 192-195, 239-242

N-myristoylation sites.

amino acids 33-38, 95-100, 116-121, 215-220, 272-277

Microbodies C-terminal targeting signal.

amino acids 315-317

Cytochrome c family heme-binding site signature.

amino acids 9-14

FIGURE 117

GAGCTCCCTCAGGAGCGCGTTAGCTTCACACCTCGGCAGCAGGAGGGCGGCAGCTCTCGCAGGCGCA
GGGCGGGCGGCAGGATCATGTCCACCACCATGCCAAGTGGTGGCGTCTCCTGTCCATCCTGGGCT
GCCGGCTGCATCGCGCCACCGGGATGGACATGTGGAGCACCAGGACCTGTACGACAACCCCGTCACCT
CCGTGTTCCAGTACGAAGGGCTCTGGAGGAGCTCGTGAGGCAGAGTCAGGCTCACCGAATGCAGGCC
TATTCACCATCCTGGACTTCCAGCCATGCTGCAGGCAGTGCAGGCCGTGATGATCGTAGGCATCGCCT
GGGTGCCATTGGCCTCCTGGTATCCATCTTGCCCTGAAATGCATCCGCATTGGCAGCATGGAGGACTCTG
CCAAAGCCAACATGACACTGACCTCCGGGATCATGTTCATTGTCTCAGGTCTTGTGCAATTGCTGGAGTG
TCTGTGTTGCCAACATGCTGGTACTAACCTCTGGATGTCCACAGCTAACATGTACACCGGCATGGGTGG
GATGGTGCAGACTGTTCAGACCAGGTACACATTGGTGCAGGCTCTGTTGCTGGCTGGCTGGCTGGAGGCC
TCACACTAACATTGGGGTGTGATGATGTGCATGCCCTGCCGGGCTGGCACAGAAGAAACCAACTACAAA
GCCGTTCTTATCATGCCCTCAGGCCACAGTGTGCCTACAAGCCTGGAGGCTCAAGGCCAGCACTGGCTT
TGGGTCCAACACCAAAAACAAGAAGATATACGATGGAGGTGCCGCACAGAGGACGAGGTACAATCTTATC
CTTCCAAGCACGACTATGTTAATGCTCTAACGACTCTCAGCACGGCGGAAGAAACTCCCAGAGCTCA
CCCCAAAAACAAGGAGATCCCCTAGATTCTTCTGCTTTGACTCACAGCTGGAAGTTAGAAAAGCCT
CGATTCATTTGGAGAGGCCAAATGGTCTTAGCCTCAGTCTCTAAATATTCCACCATAAAACA
GCTGAGTTATTATGAATTAGAGGCTATAGCTCACATTCAATCCTCTATTCTTTTAAATATAACT
TTCTACTCTGATGAGAGAATGTGGTTAACTCTCTCACATTGATGATTAGACAGACTCCCCCTC
TTCCCTCTAGTCATAAAACCCATTGATGATCTATTCCAGCTTATCCCCAGAAAACCTTGAAGGAAA
GAGTAGACCCAAAGATGTTATTCTGCTGTTGAATTGTCTCCCCACCCCAACTGGCTAGTAATAA
ACACTTACTGAAGAAGAAGCAATAAGAGAAAGATATTGTAATCTCTCCAGGCCATGATCTGGTTTCTT
ACACTGTGATCTAAAAGTTACCAAACCAAGTCATTTCAGTTGAGGCAACCAACCTTCTACTGCTG
TTGACATCTCTTATTACAGCAACACCATTCTAGGAGTTCCCTGAGCTCTCCACTGGAGTCCTTTCTGT
CGCGGGTCAGAAATTGTCCTAGATGAATGAGAAAATTATTTTTAAATTAAAGTCCTAAATATAGTTAA
AATAAATAATGTTAGTAAATGATAACTATCTCTGTGAAATAGCCTCACCCCTACATGTGGATAGAAG
GAAATGAAAAAATAATTGCTTGACATTGCTATATGGTACTTTGTAAAGTCATGCTTAAGTACAAATTCC
ATGAAAAGCTCACACCTGTAATCCTAGCACTTGGAGGCTGAGGAGGAAGGATCACTGAGGCCAGAAGT
TCGAGACTAGCCTGGCAACATGGAGAAGCCCTGTCTCACAAATACAGAGAGAAAATGCCAGTC
TGGTGGCATAACCTGTAAGTCCAGCATTCCGGAGGCTGAGGTGGAGGATCACTGAGGCCAGGGAGGT
TGGGGCTGCAGTGAGCCATGATCACACCACTGCACTCCAGCCAGGTGACATAGCGAGATCCTGTCTAAAAA
AATAAAAATAATAATGGAACACAGCAAGTCCTAGGAAGTAGGTAAAACTAATTCTTAA

FIGURE 118

MSTTCQVVAFLSILGLAGCIAATGMDMWSTQDLYDNPVTSVFQYEGLWRSCVRQSSGFTECRP
YFTILGLPAMLQAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANMTLTSGIMFIVSGL
CAIAGVSVFANMLVTNFWMSTANMYTGMGGMVQTVQTRYTFGAALFVGWVAGGLTLIGGVMCIA
CRGLAPEETNYKAVSYHASGHSVAYKPGGFCASTGFGSNTKNKKIYDGGARTEDEVQSYP SKHDY
V

Signal peptide:

amino acids 1-23

Transmembrane domains:

amino acids 81-100, 121-141, 173-194

FIGURE 119

GGAAAAACTGTTCTCTGTGGCACAGAGAACCCGCTTCAAAGCAGAAGTAGCAGTCCGGAGTCC
AGCTGGCTAAAACATCCCAGAGGATAATGGCAACCCATGCCTAGAAATCGCTGGCTGTTCTG
GTGGTGTGGAATGGTGGGACAGTGGCTGTCAGTGCATGCCTCAGTGGAGAGTGTGCGCCTTCATT
GAAAACAACATCGTGGTTTGAAAACCTCTGGGAAGGACTGTGGATGAATTGCGTGAGGCAGGCTAA
CATCAGGATGCAGTGCAAAATCTATGATTCCCTGCTGGCTCTTCTCCGGACCTACAGGCAGCCAGAG
GAATGATGTGCTGCTTCCGTGATGTCCTCTGGCTTCTGATGCCATCCTGGCATGAAATGC
ACCAGGTGCACGGGGACAATGAGAAGGTGAAGGCTCACATTCTGCTGACGGCTGGAATCATCTCAT
CATCACGGGCATGGTGGTGCTCATCCCTGAGCTGGTTGCCATGCCATCATCAGAGATTTCTATA
ACTCAATAGTGAATGTTGCCAAAAACGTGAGCTGGAGAAGCTCTACTTAGGATGGACCACGGCA
CTGGTGCTGATTGTTGGAGGAGCTGTTCTGCTGCTGCTTGTGCAACGAAAAGAGCAGTAGCTA
CAGATACTCGATACCTCCCACATCGACAACCCAAAAAGTTATCACACCGGAAAGTCACCGAGCG
TCTACTCCAGAAGTCAGTATGTGTAGTTGTATGTTTAACTTACTATAAGCCATGCAAATG
ACAAAAATCTATATTACTTCTCAAAATGGACCCAAAGAAACTTGTATTACTGTTCTAACTGCCT
AATCTTAATTACAGGAACTGTGCATCAGCTATTATGATTCTATAAGCTATTCAGCAGAATGAGATA
TTAAACCAATGCTTGATTGTTCTAGAAAGTATAGTAATTGTTCTAAGGTGGTCAAGCATCTA
CTCTTTTATCATTACTCAAAATGACATTGCTAAAGACTGCATTATTTACTACTGTAATTCTCC
ACGACATAGCATTATGTACATAGATGAGTGTAAACATTATCTCACATAGAGACATGCTTATATGGT
TTTATTAAATGAAATGCCAGTCCATTACACTGAATAAATAGAACTCAACTATTGCTTTCAGGGAA
ATCATGGATAGGGTTGAAGAAGGTTACTATTAATTGTTAAAACAGCTTAGGGATTAATGTCCTCCA
TTTATAATGAAGAGATTAAATGAAGGCTTAATCAGCATTGTAAGGAAATTGAATGGCTTCTGATAT
GCTGTTTTAGCCTAGGAGTTAGAAATCCTAACTCTTATCCTCTCCAGAGGCTTTTT
CTTGTGTATTAATTAACATTTAAAACGCAGATATTGTCAGGGCTTGCATTCAAATGCTT
TTCCAGGGCTACTCAGAAGAAAGATAAAAGTGTGATCTAAGAAAAGTGTGTTAGGAAAGTG
AAAATTTAATAAGTATTGAGTACAGACTTGAGGTTCATCAATATAAATAAAAGAGCAGAAAATA
TGTCTGGTTTCATTGCTTACCAAAAAACAAACAACAAAAAGTTGTGCTTGTGAGAACTTCACCT
GCTCCTATGTGGGTACCTGAGTCAAAATTGTCATTGTTCTGTGAAAATAAATTCTCTTGTA
CCATTCTGTTAGTTACTAAAATCTGAAATACTGTATTCTGTTATTCAAATTGATGAA
ACTGACAATCCAATTGAAAGTTGTGTCGACGTGTCTAGCTAAATGAATGTGTTCTATTGCTT
TATACATTATTAATAAAATTGTACATTTCTAATT

FIGURE 120

MATHALEIAGLFLGGVGMVGTAVTVMPQWRVSAFIENNIVVFENFWEGLWMNCVRQANIRMQCK
IYDSLLALSPDLQAARGLMCAASVMSFLAFMMAILGMKCTRCTGDNEKVKAHILLTAGIIFIITG
MVVLIPVSWVANAIIRDYNSIVNVAQKRELGEALYLGWTTALVLIVGGALFCCVFCCNEKSSSY
RYSIPSHRTTQKSYHTGKKSPSVYSRSQYV

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 82-101, 118-145, 164-188

FIGURE 121

GGAGAGAGGGCGCGCGGGTGAAAGGCGCATTGATGCAGCCTGCGGCGGCCTCGGAGCGCGGGAG
CCAGACGCTGACCACGTTCCCTCCTCGGTCTCCTCCGCCTCCAGCTCCGCCTGCCCGAGCC
GGGAGCCATGCGACCCCAGGGCCCCGCGCCTCCCCGAGCGGCTCCGCGGCCTCTGCTGCTCC
TGCTGCTGCAGCTGCCCGCGCGTCAGCGCCTCTGAGATCCCCAAGGGGAAGCAAAAGGCGCAG
CTCCGGCAGAGGGAGGTGGTGGACCTGTATAATGGAATGTGCTTACAAGGGCCAGCAGGAGTGCC
TGGTCGAGACGGGAGCCCTGGGGCAATGTTATTCCGGTACACCTGGATCCAGGTGGATG
GATTCAAAGGAGAAAAGGGGAATGCTGAGGGAAAGCTTGAGGAGTCCTGGACACCCAAC
AAGCAGTGTTCATGGAGTTCAATTGAATTATGGCATAGATCTGGAAAATTGCGGAGTGTACATT
TACAAAGATGCGTTCAAATAGTGCTCTAAGAGTTGTTCACTGGCTCACTCGGCTAAAATGCA
GAAATGCATGCTGTCAGCGTTGGTATTCACATTCAATGGAGCTGAATGTTCAAGGACCTCTCCC
ATTGAAGCTATAATTATTGGACCAAGGAAGCCCTGAAATGAATTCAACAAATTATTCATCG
CACTCTTCTGTGGAAGGACTTGTGAAGGAATTGGTGCTGGATTAGTGGATGTTGCTATCTGG
TTGGCACTTGTTCAGATTACCCAAAAGGAGATGTTCACTGGATGGAATTCAAGTTCTCGCATC
ATTATTGAAGAACTACCAAAATGCTTAATTTCATTGCTACCTCTTTTATTATGCC
TTGGAATGGTCACTTAAATGACATTAAATAAGTTATGTATACATCTGAATGAAAAGCAAAG
CTAAATATGTTACAGACCAAGTGTGATTCACACTGTTAAATCTAGCATTTCATTG
CTTCAATCAAAAGTGGTTCAATATTTTTAGTTGGTTAGAATACTTCTTCATAGTCACATT
CTCTCAACCTATAATTGGAATTGTTGGTCTTTGTTCTTAGTATAGCATTGTTA
AAAAAAATAAAAAGCTACCAATCTTGTACAATTGAAATGTTAAGAATTTTTATCTGT
TAAATAAAATTATTCCAACA

FIGURE 122

MRPQGPAA SPQR LRG LLLL LQLPAPSSASEIPKGKQKAQLRQREVVDLYNGMCLQGPAGVPGR
DGSPGANVIPGTPGIPGRDGFKGEKGECLRESFEESWTPNYKQCSWSSLNYGIDLGKIAECTFTK
MRSNSALRVLFSGSLRLKCRNACCQRWYFTFNGAECSGPLPIEAIYLDQGSPEMNSTINIHRTS
SVEGLCEGIGAGLVDVAIWVGTCSDYPKGDASTGWNSVSRIIEELPK

Signal peptide:

amino acids 1-30

Transmembrane domain:

amino acids 195-217

FIGURE 123

GCTGAGCGTGTGCGCGGTACGGGCTCTCCTGCCTCTGGGCTCCAACGCAGCTCTGGCTGAA
CTGGGTGCTCATCACGGGAAC TGCTGGCTATGAAATACAGATGTGGCAGCTAGGTAGCCCCAA
ATTGCCTGGAAGAATACATCATGTTTCGATAAGAAGAAATTGAGGATCCAGTTTTTTTA
ACCGCCCCCTCCCCACCCCCCAAAAAACTGTAAGAGATGCAAAACGTAATATCCATGAAGATCC
TATTACCTAGGAAGATTGATGTTGCTGCGAATGCAGTGGGATTATTGTCTTGGAG
TGTTCTGCGTGGCTGGCAAAGAATAATGTCAAAATCGTCCATCTCCAAGGGTCCAATT
TCTTCTGGGTGTCAGCGAGCCCTGACTCACTACAGTGCAGCTGACAGGGCTGTCACTG
GCCCTAAGCCAAGCAGACCTAAGGACGACCTTGAACAATACAAAGGATGGTTCAATG
TAATTAGGCTACTGAGCGGATCAGCTGTAGCACTGGTTAGCCCCACTGTCTTACTGACAATG
CTTTCTTCTGCCGAACGAGGATGCCCTAAGGGCTGTAGGTGAAGGAAATGGTATATTGTGA
ATCTCAGAAATTACAGGAGATACCCCTAAGTATATCTGCTGGTGCTTAGGTTGTCCCTCGCT
ATAACAGCCTCAAAAACCTTAAGTATAATCAATTAAAGGGCTCAACCAGCTCACCTGGCTATAC
CTTGACCATAACCATACTCAGCAATTGACGAAATGCTTTAATGGAATACGAGACTCAAAGA
GCTGATTCTTAGTTCCAATAGAATCTCCTATTTCTTAACAATACCTCAGACCTGTGACAAATT
TACGGAACTTGGATCTGCTTATAATCAGCTGCATTCTCTGGGATCTGAACAGTTGGGGCTTG
CGGAAGCTGCTGAGTTACATTACGGCTAACCTCCCTGAGAACCATCCCTGTGCGAATTCCA
AGACTGCCGCAACCTGGAACCTTGGACCTGGATATAACCGGATCCGAAGTTAGCCAGGAATG
TCTTGCTGGCATGACTCAAAGAACCTCACCTGGAGCACAATCAATTCCAAGCTCAAC
CTGGCCCTTTCCAAGGGTGGTCAGCCTCAGAACCTTACTTGCAGTGGAAATAAATCAGTGT
CATAGGACAGACCATGTCCTGGACCTGGAGCTCCTACAAAGGCTGATTATCAGGAATGAGA
TCGAAGCTTCAGTGGACCCAGTGTGTTCCAGTGTGCTCCGAATCTGCAGCGCCTAACCTGGAT
TCCAACAAGCTCACATTATTGGTCAAGAGATTGGATCTTGGATATCCCTCAATGACATCAG
TCTTGCTGGGAATATGGGAATGCAAGCAGAAATATTGCTCCCTGTAAACTGGCTGAAAAGTT
TTAAAGGTCTAAGGGAGAATACAATTATCTGTGCCAGTCCAAAGAGCTGCAAGGAGTAAATGTG
ATCGATGCAGTGAAGAACTACAGCATCTGGCAAAAGTACTACAGAGAGGTTGATCTGCCAG
GGCTCTCCCAAAGCCGACGTTAACGCCCAGCTCCCCAGGCCGAAGCATGAGAGCAAACCCCTT
TGCCCCCGACGGTGGAGCCACAGAGCCGGCCAGAGACCGATGCTGACGCCGAGCACATCT
TTCCATAAAATCATCGCGGGCAGCGTGGCGTTTCTGTCCGTGCTCGTCATCCTGCTGGTTAT
CTACGTGTCATGGAAGCGGTACCCCTGCGAGCATGAAGCAGCTGCAGCAGCGCTCCCTCATGCGAA
GGCACAGGAAAAAGAAAAGACAGTCCCTAAAGCAAATGACTCCCAGCACCCAGGAATTATGTA
GATTATAAACCCACCAACACGGAGACCAGCGAGATGCTGCTGAATGGACGGGACCCCTGCACCTA
TAACAAATCGGGCTCCAGGGAGTGTGAGGTATGAACCATGTAAGAGCTCTAAAAGCT
GGGAAATAAGTGGTGTCTTATTGAACTCTGGTACTATCAAGGGAACCGCGATGCCCTCC
TTCCCTCTCCCTCTCACTTGGTGGCAAGATCCTCCTGTCCGTTAGTGCATTATAACT
GGTCAATTCTCTCATACATAATCAACCCATTGAAATTAAATACCACAATCAATGTGAAGCTT
GAACTCCGGTTAATATAATACCTATTGTATAAGACCCCTTACTGATTCCATTAAATGTCGCAATT
GTTTAAGATAAAACTCTTCATAGGTAAAAAAAAAA

FIGURE 124

MGFNVIRLLSGSAVALVIAPTVLLTMLSSAERGCPKGRCRCEGKMVYCESQKLQEIPSSISAGCLG
LSLRYNSLQKLKYNQFKGLNQLTWLYLDHNHISNIDENAFNGIRRLKELILSSNRISYFLNNTR
PVTNLRNLDLSYNQLHSLGSEQFRGLRKLLSLHLRSNSLRTIPVRIFQDCRNLELLDLGYNRIRS
LARNVFAGMIRLKELHLEHNQFSKLNLAFLPRLVSLQNLYLQWNKISVIGQTMSTWSSLQRLDL
SGNEIEAFSGPSVFQCPNLQRLNLDNKLTFIGQEILDWSWISLNDISLAGNIWECSRNICSIVN
WLKSFKGLRENTIICASPKELQGVNVIDAVKNYSICGKSTTERFDLARALPKPTFKPKLPRPKHE
SKPPLPPTVGATEPGPETDADAEHISFHKIIAGSVALFLSVLVIILVIYVSWKRYPASMKQLQQR
SLMRRHRKKRQSLKQMTPSTQE FYVDYKPTNTETSEMLLNGTGPCYNKSGSRECEV

Important features of the protein:

Signal peptide:

amino acids 1-33

Transmembrane domain:

amino acids 420-442

N-glycosylation sites.

amino acids 126-129, 357-360, 496-499, 504-507

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 465-468

Tyrosine kinase phosphorylation site.

amino acids 136-142

N-myristylation sites.

amino acids 11-16, 33-38, 245-250, 332-337, 497-502, 507-512

FIGURE 125

CCGTTATCGTCTTGCCTACTGCTGAATGTCCGTCCGGAGGAGGGAGAGGCTTTGCCGCTG
ACCCAGAGATGGCCCCGAGCGAGCAAATTCTACTGTCCGGCTGCCGGCTACCGTGGCCGAGCT
AGCAACCTTCCCCTGGATCTCACAAAAACTCGACTCCAAATGCAAGGAGAAGCAGCTTGCCTC
GGTGGGAGACGGTGCAAGAGAATCTGCCCTATAGGGGAATGGTGCGCACAGCCCTAGGGATC
ATTGAAGAGGAAGGCTTCTAAAGCTTGGCAAGGAGTGACACCCGCCATTACAGACACGTAGT
GTATTCTGGAGGTCGAATGGTCACATATGAACATCTCGAGAGGTTGTGTTGGCAAAAGTGAAG
ATGAGCATTATCCCCTTGGAAATCAGTCATTGGAGGGATGATGGCTGGTGTATTGCCAGTT
TTAGCCAATCCAACGTACCTAGTGAAGGTTAGATGCAAATGGAAGGAAAAGGAAACTGGAAGG
AAAACCATTGCGATTCGTGGTGTACATCATGCATTGCAAAATCTTAGCTGAAGGAGGAATAC
GAGGGCTTGGCAGGCTGGTACCCAATATAAAAGAGCAGCAGTGGTGAATATGGGAGATTAA
ACCACTTATGATAACAGTGAACACTACTTGGTATTGAATACACCACTTGAGGACAATATCATGAC
TCACGGTTATCAAGTTATGTTCTGGACTGGTAGCTCTATTCTGGAACACCAGCCATGTCA
TCAAAAGCAGAATAATGAATCAACCACGGATAAAACAAGGAAGGGACTTTGTATAAATCATCG
ACTGACTGCTTGATTCAAGGTGAAGGATTGAGTCTATATAAAGGCTTTACC
ATCTGGCTGAGAATGACCCCTGGTCAATGGTGTCTGGCTTACCTATGAAAAAAATCAGAGAGA
TGAGTGGAGTCAGTCATTAA

FIGURE 126

MSVPEEEERLLPLTQRWPRASKFLLSGCAATVAELATFPLDLTKTRLQMQGEAALARLGDGARES
APYRGMVRTALGIIEEEGFLKLWQGVTPAIYRHVVYSGGRMVTYEHLCVVFGKSEDEHYPLWKS
VIGGMMAGVIGQFLANPTDLVKVQMQMEGKRKLEGKPLRFRGVHAFAKILAEGGIRGLWAGWVP
NIQRAALVNMGDLTYDTVKHYLVLNTPLEDNIMTHGLSSLCGLVASILGTPADVIKSRI MNQP
RDKQGRGLLYKSSTDCLIQAVQGEGFMSLYKGFLPSWLRMTPWSMVFWLTYEKIREMSGVSPF

Transmembrane domains:

amino acids 25-38, 130-147, 233-248

FIGURE 127

CGCGGATCGGACCCAAGCAGGTGGCGGCCGGCAGGAGAGCAGGCCGGCGTCAGCTCCTCGAC
CCCCGTGTCGGCTAGTCCAGCGAGGCGGACGGCGCGTGGGCCATGCCAGGCCGGCATGG
AGCGGTGGCGCGACCGGCTGGCGCTGGTACGGGGCCTCGGGGGCATCGCGCGGCCGTGGCC
CGGGCCCTGGTCCAGCAGGGACTGAAGGTGGTGGCTGCGCCCGCACTGTGGCAACATCGAGGA
GCTGGCTGCTGAATGTAAGAGTGCAGGCTACCCGGGACTTGTACAGATGTGACCTAT
CAAATGAAGAGGACATCCTCTCCATGTTCTCAGCTATCCGTTCTCAGCACAGCGGTGTAGACATC
TGCATCAACAATGCTGGCTGGCCCGCCCTGACACCCCTGCTCAGGCAGCACAGCGGTGTAGACATC
GGACATGTTCAATGTGAACGTGCTGGCCCTCAGCATCTGCACACGGGAAGCCTACCAGTCCATGA
AGGAGCGGAATGTGGACGATGGGACATCATTAAACATCAATAGCATGTCTGGCCACCGAGTGTAA
CCCCTGTCTGTGACCCACTTCTATAGTGCCACCAAGTATGCCGTCACTGCGCTGACAGAGGGACT
GAGGCAAGAGCTCGGGAGGCCAGACCCACATCCGAGCCACGTGCATCTCTCAGGTGTGGTGG
AGACACAATTGCCTCAAACCCGAGGATGTGCCGAGGCTGTTATCTACGTCTCAGCACCCCCGACA
ATGAAGTGTCTCAAACCCGAGGATGTGCCGAGGCTGTTATCTACGTCTCAGCACCCCCGACA
CATCCAGATTGGAGACATCCAGATGAGGCCACGGAGCAGGTGACACTGTGACTGTGGGAGCTCC
TCCTTCCCTCCCCACCCCTCATGGCTGCCTGCCTGGATTAGGTGTTGATTCTGGAT
CACGGGATACCACCTCCTGTCCACACCCGACCAGGGCTAGAAAATTGTTGAGATTTATA
TCATCTTGTCAAATTGCTTCAGTTGTAATGTGAAAAATGGCTGGGAAAGGAGGTGGTGTCCC
TAATTGTTTACTTGTAACTTGTCTGTGCCCTGGCACTTGGCCTTGTCTGCTCTCAGTG
TCTTCCCTTGACATGGAAAGGAGTTGTGCCAAAATCCCCATCTTGTGACCTCAACGTCTG
TGGCTCAGGGCTGGGTGGCAGAGGGAGGCCTCACCTATATCTGTGTTATCCAGGGCTCC
AGACTTCCTCCTGCCTGCCCACTGCACCCCTCCCCCTATCTCCTCTCGGCTCCCC
AGCCCAGTCTGGCTTGTCCCCTGGGTCACTCCACTCTGACTCTGACTATGGCAG
CAGAACACCAGGGCTGGCCAGTGGATTTCATGGTGATCATTAAAAAGAAAATCGAACCAA
AAAAAAA

FIGURE 128

MARPGMERWRDRRLALVTGASGGIGAAVARALVQQGLKVVGCARTVGNIEELAAECKSAGYPGTLI
PYRCDLSNEEDILSMFSAIRSQHSGVDICINNAGLARPDTLLSGSTSGWKDMFNVLALSICTR
EAYQSMKERNVDDGHIININMSGHRVLPLSVTHFYSATKYAVTALTEGLRQELREAQTHIRATC
ISPGVVETQFAFKLHDKDPEKAAATYEQMKCLKPEDVAEAVIYVLSTPAHIQIGDIQMRPTEQVT

Important features of the protein:

Signal peptide:

amino acids 1-17

N-myristoylation sites.

amino acids 18-24, 21-27, 22-28, 24-30, 40-46, 90-96, 109-115,
199-205

Short-chain alcohol dehydrogenase.

amino acids 30-42, 104-114

FIGURE 129

AACTCTACATGGGCCTCCTGCTGGTGCCTTCCTCAGCCTCCTGCCGGTGGCCTACACCAT
CATGTCCTCCCACCCTCCTTGACTGCGGGCCGTTCAAGGTGCAGAGTCTCAGTTGCCGGAGC
ACCTCCCCTCCCGAGGCAGTCTGCTCAGAGGGCCTCGGCCAGAATTCCAGTTCTGGTTCATGC
CAGCCTGTAAAAGGCCATGGAACTTGGGTGAATCACCGATGCCATTAAAGAGGGTTCTGCCA
GGATGGAAATGTTAGGTCGTTCTGTCTGCCTGTTCAATTCAAGTAGCCACCAGCACCTGTGG
CCGTTGAGTGCTTGAAATGAGGAACTGAGAAAATTAAATTCTCATGTATTTCTCAATTATTAA
TTAATTAACTGATAGTTGTACATATTGGGGTACATGTGATATTGGATACTGTATACAA
TATATAATGATCAAATCAGGGTAACGGGATATCCATCACATCAAACATTATTTTATTCTTT
TTAGACAGAGTCTCACTCTGTCACCCAGGCTGGAGTGCAGTGGTGCATCTCAGTTACTGCAAC
CTCTGCCTGCCAGGTTCAAGCGATTCTCATGCCCTCACCTCCAAAGTAGCTGGACTACAGGCAT
GCACCACAAATGCCCAACTAATTGTATTTAGTAGAGACGGGTTTGCATGTGCCAGG
CTGGCCTGAACCTGGCCTCAAACAATCCACTGCCTGGCCTCCAAAGTGTATTGATTACA
GGCGTGAGCCACCGTGCCTGGCTAAACATTATCTTTCTTGTTGGAACTTGAAATTAT
ACAATGAATTATTGTTAACTGTCATCTCCCTGCTGTGCTATGGAACACTGGGACTTCTCCCTCT
ATCTAACTGTATATTGTACCACTAACCGTACTTCATCCCCACTCCTCTATCCTTCCC
AACCTCTGATCACCTCATTCTACTCTACCTCCATGAGATCCACTTTAGCTCCACATGTG
AGTAAGAAAATGCAATATTGTCTTCTGTGCCTGGTTATTCACCTAACATAATGACTCCTG
TTCCATCCATGTTGCTGCAAATGACAGGATTGTTCTTAATTCAATTAAAATAACCACACATG
GCAAAAAA

FIGURE 130

MGLLLLVLFLSLLPVAYTIMITSLPPSFDCGPFRCRVSVAREHLPGRGSLLRGPRPRIPLVSCQPV
KGHGTLGESPMPFKRVFCQDGNVRSFCVCAVFSSHQPPVAVECLK

Important features of the protein:

Signal peptide:

amino acids 1-18

N-myristoylation site.

amino acids 86-92

Zinc carboxypeptidases, zinc-binding region 2 signature.

amino acids 68-79

FIGURE 131

TTCTGAAGTAACGGAAGCTACCTGTATAAAGACCTAACACTGCTGACCATGATCAGCGCAGCCTGGAGC
ATCTCCTCATCGGGACTAAAATTGGGCTGTCCTCAAGTAGCACCTCTATCAGTTATGGCTAAATCCTG
TCCATCTGTGTGCGCTGCATGCGGTTTCACTGTAATGATCGCTTCTGACATCCATTCCAACAG
GAATACCAGAGGATGCTACAACCTCTACCTTCAGAACAAACAAATAATGCTGGATTCCCTCAGAT
TTGAAAAACTGCTGAAAGTAGAAAGAATATACCTATACCACAAACAGTTAGATGAATTCCCTACCAACCT
CCCAAAGTATGTAAAAGAGTTACATTGCAAGAAAATAACATAAGGACTATCACTTATGATTCACTTCAA
AAATTCCCTATCTGGAAGAATTACATTAGATGACAACACTGTCTCTGCAGTTAGCATAGAAGAGGGAGCA
TTCCGAGACAGCAACTATCTCGACTGCTTTCTGCCCCTGAATCACCTAGCACAATTCCCTGGGTTT
GCCAGGACTATAGAAGAACTACGCTTGGATGATAATCGCATATCCACTATTCATCACCCTCTTCAAG
GTCTCACTAGTCTAAAACGCCTGGTCTAGATGGAAACCTGTTGAACAATCATGGTTAGGTGACAAAGTT
TTCTCAACCTAGTTAATTGACAGAGCTGTCCTGGTGCAGATTCCCTGACTGCTGCACCAGTAAACCT
TCCAGGCACAAACCTGAGGAAGCTTATCTCAAGATAACCACATCAATCGGTCCCCCAAATGCTTTT
CTTATCTAAGGCAGCTATCGACTGGATATGTCCAATAAACCTAACCTAACGTTACGGGTATCTT
GATGATTGGACAATATAACACAACGTTACCTACCTGTGAAGGTCAACGTGCGTGGCTCATGTGCCAAGCCCCAGAAA
AGGTTCGTGGATGGCTATTAAGGATCTCAATGCAGAACTGTTGATTGTAAGGACAGTGGGATTGTAAGC
ACCATTCAAGATAACCACGTCAATACCCAACACAGTGTATCCTGCCAAGGACAGTGGCCAGCTCCAGTGAC
CAAACAGCCAGATATTAAGAACCCCAAGCTCACTAAGGATCAACAAACCACAGGGAGTCCCTCAAGAAAAAA
CAATTACAATTACTGTGAAGTCTGTACACCTCTGATACCATTCAATCTCTGGAAACTTGCTCTACCTATG
ACTGCTTGAGACTCAGCTGGCTAAACTGGGCCATAGCCCGCATTGGATCTATAACAGAAACAATTGT
AACAGGGAACCGCAGTGAAGTACTGGTCACAGCCCTGGAGCCTGATTCAACCTATAAAAGTATGCATGGTTC
CCATGGAAACCGAACCTCTACCTATTGATGAAACTCCCTGTTGTATTGAGACTGAAACTGCACCCCTT
CGAATGTACAACCCCTACAACCACCCCTCAATCGAGAGCAAGAGAGAACCTTACAAAAACCCCAATTAC
TTGGCTGCCATCATTGGTGGGCTGTGGCCCTGGTTACCATGCCCTCTTGCTTAGTGTGTTGGTATG
TTCATAGGAATGGATCGCTTCTCAAGGAACCTGTGCATATAGCAAAGGGAGGAGAACGGATGACTAT
GCAGAAGCTGGCACTAAGAAGGACAACCTATCCTGGAAATCAGGGAAACTTCTTCAGATGTTACCAAT
AAGCAATGAACCCATCTCGAAGGAGGAGTTGTAATACACACCATATTCCCTCTAATGGAATGAATCTGT
ACAAAAACAATCACAGTGAAAGCAGTAGTAACCGAAGCTACAGAGACAGTGGTATTCCAGACTCAGATCAC
TCACACTCATGATGCTGAAGGACTCACAGCAGACTTGTGTTGGTTTTAACCTAAGGGAGGTGATG
GT

FIGURE 132

MISAAWSIFLIGTKIGLFLQVAPLSVMAKSCPSVRC DAGFIYCNDRFLTSIPTGIPEDATTLYL
QNNQINNAGIPSDLKNLLKVERIYLYHNSLDEFPTNLPKYVKELHLQENNIRTITYDSL_SKIPYL
EELHLDDNSVSAVSIEEGAFRDSNYLRLLFLSRNHLSTIPWGLPRTIEELRLDDNRISTISSPSL
QGLTSLKRLVLDGNLLNNHGLDKVFFNLVNLTELSLVRNSLTAAPVNLPGTNLRKLYLQDNHIN
RVPPNAFSYLRQLYRLDMSNNNLSNLPQGIFDDLDNITQLILRNNPWYCGCKMKWVRDWLQSLPV
KVNVRGLMCQAPEKVRGMAIKDLNAELFDCKDSGIVSTIQTITTAIPNTVYPAQGQWPAPVTKQPD
IKNPKLTKDQQTGSPSRKTITITVKS VTS DTIHISWKLALPMTALRLSWLKGHSPAFGSITET
IVTGERSEYLVTALEPDSPYKVCVPMETS NLYLFDETPVCIETETAPLRMYNPTTLNREQEKE
PYKNPNLPLAIIIGGAVALVTIALLALVCWYVHRNGSLFSRNCAYSKGRRRKDDYAEAGTKKD NS
ILEIRETSFQMLPISNEPISKEEFVIHTIFPPNGMNLYKNNHSESSSNRSYRDSGIPDSDHSHS

Important features of the protein:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 531-552

N-glycosylation sites.

amino acids 226-229, 282-285, 296-299, 555-558, 626-629, 633-636

Tyrosine kinase phosphorylation site.

amino acids 515-522

N-myristoylation sites.

amino acids 12-17, 172-177, 208-213, 359-364, 534-539, 556-561,
640-645

Amidation site.

amino acids 567-570

Leucine zipper pattern.

amino acids 159-180

Phospholipase A2 aspartic acid active site.

amino acids 34-44

FIGURE 133

CCGTCATCCCCCTGCAGCCACCCCTCCAGACTCTTGCCTCAGGCCACCCAGGCTCTGGCA
GCCCTGCCGGCCACTTGTCTTCCATGTCTGCCAGGGGAGGTGGGAAGGAGGTGGGAGGGCG
TGCAGAGGCAGTCTGGCTTGGCCAGAGCTCAGGGTGTGACCGTGTGACCAAGCAGTGACAGAG
GCCGCCATGCCAGCCTGGGCTGCTGCTCTGCTTACTGACAGCACTGCCACCGCTGTGGT
CCTCCTCACTGCCCTGGGCTGGACACTGCTGAAAGTAAAGCCACCATTGCAGACCTGATCCTGTCT
GCGCTGGAGAGAGGCCACCGTCTCCTAGAACAGAGGCTGCCGTGAAATCAACCTGGATGGCATGGT
GGGGGTCCGAGTGCTGGAAGAGCAGCTAAAAGTGTCCGGGAGAAGTGGGCCAGGAGCCCTGC
TGCAGCCGCTGAGCCTGCGCGTGGGATGCTGGGGAGAAGCTGGAGGCTGCCATCCAGAGATCC
CTCCACTACCTCAAGCTGAGTGATCCAAAGTACCTAACAGAGAGTCCAGCTGACCCCTCAGCCCG
GTTTGGAAAGCTCCCACATGCCCTGGATCCACACTGATGCCCTTGGTGTACCCACGTTGGC
CCCAGGACTCATTCTCAGAGGAGAGAAGTGACGTGTGCCCTGGTGCAGCTGCTGGAAACCGGGACG
GACAGCAGCGAGCCCTGCCCTCAGACCTCTGCAGGAGCCTCATGACCAAGCCGGCTGCTC
AGGCTACTGCCGTCCCACCAACTGCTCTTCTCTGGCCAGAATGAGGGGATGCACACAGG
GACCACTCCAACAGAGCCAGGACTATATCAACCTCTGCGCAACATGATGGACTTGAACCGC
AGAGCTGAGGCCATCGGATACGCCCTACCCCTACCCGGACATCTCATGGAAAACATCATGTTCTG
TGGAATGGCGGCTCTCGACTTACAAGCTCCGGTGGCTGGAGGCCATTCTCAGCTGGCAGA
AACAGCAGGAAGGATGCTTCGGGAGCCTGATGCTGAAGATGAAGAATTATCTAAAGCTATTCAA
TATCAGCAGCATTTCGAGGAGAGTGAAGAGGCAGAAAAACAATTCCAGATTCTCGCTCTGT
TGCTCAGGCTGGAGTACAGTGGCGCAATCTGGCTCACTGCAACCTTGCCTCTGGTTCAAGC
AATTCTCTGCCCTCATCCTCCGAGTAGCTGGACTACAGGAGCGGCCACCATACCTGGCTAAT
TTTATATTTTTAGTAGAGACAGGGTTCATCATGTTGCTCATGCTGGTCTCGAACTCCTGAT
CTCAAGAGATCCGCCACCTCAGGCTCCAAAGTGTGGATTATAGGTGTGAGCCACCGTGTCTG
GCTGAAAAGCACTTCAAAGAGACTGTGTTGAATAAGGGCAAGGTTCTGCCACCCAGCACTC
ATGGGGCTCTCTCCCTAGATGGCTGCTCCTCCCACAAACACAGCCACAGCAGTGGCAGCCCTGG
GTGGCTTCTATACATCCTGGCAGAATACCCCCAGCAAACAGAGAGGCCACCCATCCACACCG
CCACCACCAAGCAGCCGCTGAGACGGACGGTTCCATGCCAGCTGCCCTGGAGGGAGAACAGACCC
TTAGTCCTCATCCCTAGATCCTGGAGGGCACGGATCACATCCTGGGAAGAAGGCATCTGGAGG
ATAAGCAAAGCCACCCGACACCCAACTTGGAAAGCCCTGAGTAGGCAGGGCAGGGTAGGTGGG
GGCCGGGAGGGACCCAGGTGTGAACGGATGAATAAGTCAACTGCAACTGAAACTGAAAAAA

FIGURE 134

MSARGRWEGGRRACRGLGLARAQGAERVTSSEQRPMASLGLLLLLTLPPPLWSSSLPGLD
TAESKATIADLILSALERATVFLEQRLPEINLDGMVGVRVLEEQLKSVREKWAQEPLLQPLSLRV
GMLGEKLEAAIQRSILHYLKLSDPKYLREFQLTLQPGFWKLPHAWIHTDASLVYPTFGPQDSFSEE
RSDVCLVQLLGTGDSSEPCGLSDLRSLMTKPGCSGYCLSHQLLFFLWARMRGCTQGPLQQSD
YINLFCANMMDLNRRAEAIGYAYPTRDIFMENIMFCGMGGFSDFYKLRWLEAILSWQKQQEGCFG
EPDAEDEELSKAIQYQQHFSRRVKRREKQFPDSRSVAQAGVQWRNLGSLQPLPPGFQFSCLILP
SSWDYRSVPYLANFYIFLVETGFHHVAHAGLELLISRDPPTSGSQVGL

Important features of the protein:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 39-56

Tyrosine kinase phosphorylation sites.

amino acids 149-156, 274-282

N-myristoylation sites.

amino acids 10-16, 20-26, 63-69, 208-214

Amidation site.

amino acids 10-14

Glycoprotein hormones beta chain signature 1.

amino acids 230-237

FIGURE 135

GGTCTGAGTGCAGAGCTGCTGTCATGGCGGCCCTCTGTGGGCTCTTCCGCTGCTGCTG
CTGCTGCTATCGGGGATGTCCAGAGCTCGGAGGTGCCGGGCTGCTGCTGAGGGATCGGAGG
GAGTGGGTGGCATAGGAGATCGCTCAAGATTGAGGGCGTGCAGTTGTCAGGGTGAAGC
CTCAGGACTGGATCTGGCGGCCGAGTGCTGGTAGACGGAGAAGAGCACGTCGGTTCTTAAG
ACAGATGGGAGTTTGTGGTCATGATATACTCTGGATCTTATGTAGTGGAAAGTTGTATCTCC
AGCTTACAGATTTGATCCCGTTGAGTGGATATCACTCGAAAGGAAAATGAGAGCAAGATATG
TGAATTACATCAAAACATCAGAGGTTGTCAGACTGCCCTATCCTCTCAAATGAAATCTCAGGT
CCACCTTCTTACTTATTAAAAGGAATCGTGGGCTGGACAGACTTCTAATGAACCCAATGGT
TATGATGATGGTCTTCCTTATTGATATTGTGCTTCTGCCTAAAGTGGTCAACACAAGTGATC
CTGACATGAGACGGAAATGGAGCAGTCATGAATATGCTGAATTCCAACCAGTGGCTGAT
GTTTCTGAGTTCATGACAAGACTCTCTCTTCAAATCATCTGGCAAATCTAGCAGGGCAGCAG
TAAAACAGGCAAAAGTGGGCTGGCAAAAGGAGGTAGTCAGGCCGCCAGAGCTGGCATTGCAC
AAACACGGCAACACTGGGTGGCATCCAAGTCTGGAAAACCAGTGTGAAGCAACTACTATAAACTT
GAGTCATCCGACGTTGATCTCTTACAACGTGTATGTT
AACTTTTAGCACATGTTGTACTTGGTACACGAGAAAACCCAGCTTCATCTTGTCTGTAT
GAGGTCAATATTGATGTCAGTGAATTACAGTGTCTATAGAAAATGCCATTAATAAAATTAT
ATGAACTACTATACATTATGTATTAATTAAAACATCTTAATCCAGAAATCAAAAAAAA
AAAAAAA

FIGURE 136

MAAALWGFFPVLLLLLSDVQSSEVPGAAAEGSGGSGVGIGDRFKIEGRAVPGVKPQDWISAA
RVLVDGEEHVGLKTDGSFVVHDIPSGSYVVEVVSAYRFDPVRVDITSKGKMRARYVNYIKTSE
VVRLPYPLQMKSSGPPSYFIKRESWGWTDFLMNPVMMVLPLLIFVLLPKVVNTSDPDMRREME
QSMNMLNSNHELPDVSEFMTRLFSSKSSGKSSGSSKTGSGAGKRR

Important features of the protein:

Signal sequence:

amino acids 1-23

Transmembrane domain:

amino acids 161-182

N-glycosylation site.

amino acids 184-187

Glycosaminoglycan attachment sites.

amino acids 37-40, 236-239

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 151-154

N-myristoylation sites.

amino acids 33-38, 36-41, 38-44, 229-234

Amidation site.

amino acids 238-241

ATP/GTP-binding site motif A (P-loop).

amino acids 229-236

FIGURE 137

GATGGCGCAGCCACAGCTCTGTGAGATTGATTCTCCCCAGTTCCCTGTGGGTCTGAGGGGA
CCAGAAGGGTGAGCTACGTTGGCTTCTGGAAGGGAGGCTATATGCGTCAATTCCCCAAACAA
GTTTGACATTCCCCTGAAATGTCATTCTATCTATTCACTGCAAGTGCTGCTGTTCCAGGC
CTTACCTGCTGGGCACTAACGGCGGAGCCAGGATGGGGACAGAATAAAGGAGCCACGACCTGTGC
CACCAACTCGCACTCAGACTCTGAACCTCAGACCTGAAATCTTCTCTCACGGGAGGCTTGGCAGT
TTTCTTACTCCTGTGGTCTCCAGATTCAGGCCTAAGATGAAAGCCTCTAGTCTTGCCCTCAGC
CTTCTCTGCTGCGTTTATCTCCTATGGACTCCTCCACTGGACTGAAGACACTCAATTGGG
AAGCTGTGTGATGCCACAAACCTCAGGAAATACGAAATGGATTCTGAGATAACGGGCAGTG
TGCAAGCCAAGATGGAAACATTGACATCAGAAATCTTAAGGAGGACTGAGTCTTGCAAGACACA
AAGCCTGCGAATCGATGCTGCCCTGCCATTGCTAAGACTCTATCTGGACAGGGTATTAA
AAACTACCAGACCCCTGACCATTATACTCTCCGAAGATCAGCAGCCTGCCAATTCTTCTTA
CCATCAAGAAGGACCTCCGGCTCTCATGCCACATGACATGCCATTGTGGGGAGGAAGCAATG
AAGAAAATACAGCCAGATTCTGAGTCACTTGAAAAGCTGGAACCTCAGGCAGCAGTTGTGAAGGC
TTTGGGGGAACTAGACATTCTCTGCAATGGATGGAGGAGACAGAATAGGAGGAAAGTGTGCTG
CTGCTAAGAATATTGAGGTCAAGAGCTCCAGTCTCAATACCTGCAGAGGAGGCATGACCCCAA
ACCACCATCTTTACTGTACTAGTCTGTGCTGGTCACAGTGTATCTTATTGCTACTTG
CTTCCTGCTGATGATTGTCTTATGCATCCCCAATCTTAATTGAGACCATACTGTATAAGATT
TGTAATATCTTCTGCTATTGGATATATTATTAGTTAATATATTATTATTATTGCTATT
ATGTATTATTATTACTTGGACATGAAACTTAAAAAAATTACAGATTATTTATAACCTG
ACTAGAGCAGGTGATGTATTATACAGTAAAAAAACCTGTAAATTCTAGAAGAGTGG
CTAGGGGGTTATTCAACTAAGGACATATTACTCATGCTGATGCTGTGAGAT
ATTGAAATTGAACCAATGACTACTTAGGATGGTTGTGGAATAAGTTGATGTGGAATTGCAC
ATCTACCTTACAATTACTGACCACCCAGTAGACTCCCCAGTCCATAATTGTGTATCTCCAG
CCAGGAATCCTACACGGCCAGCATGTATTCTACAAATAAGTTCTTGCATACCAAAAAAA
AAAAAA

FIGURE 138

MRQFPKTSFDISPEMSFSIYSLQVPAVGLTCWALTAEPGWGQNKGATTCATNSHSDSELRPEIF
SSREAWQFFLLLWSPDFRPKMKASSLAFSLLSAAFYLLWTPSTGLKTLNLGSCVIATNLQEIRNG
FSEIRGSVQAKDGNIDIRILRRTESLQDTKPANRCCLLRHLLRLYLDRVFKNYQTPDHYTLRKIS
SLANSFLTIKKDLRLSHAHMTCHCGEEAMKKYSQILSHFEKLEPQAAVVKALGELDILLQWMEET
E

Important features of the protein:

Signal peptide:

amino acids 1-42

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 192-195, 225-228

N-myristoylation sites.

amino acids 42-47, 46-51, 136-141

FIGURE 139

CCTGGAGCCGGAAGCGCGGCTGCAGCAGGGCGAGGCTCCAGGTGGGTCGGTCCGCATCCAGCC
TAGCGTGTCCACGATGCGGCTGGCTCCGGACTTCGCTACCTGTTGCGTAGCGATCGAGGTGC
TAGGGATCGCGGTCTCCTCGGGATTCTTCCCGCTCCCGTCTGCCAGAGCGGAA
CACGGAGCGGAGCCCCAGCGCCGAACCCCTCGCTGGAGCCAGTCTAACTGGACCACGCTGCC
ACCACCTCTTCAGTAAAGTTGTTATTGTTCTGATAGATGCCTTGAGAGATGATTGTGTTG
GGTCAAAGGGTGTGAAATTATGCCCTACACAACCTTACCTGTGGAAAAAGGAGCATCTCACAGT
TTTGTGGCTGAAGCAAAGCCACCTACAGTTACTATGCCTCGAATCAAGGCATTGATGACGGGGAG
CCTTCCTGGCTTGTGACGTATCAGGAACCTCAATTCTCCTGCACTGCTGGAAAGACAGTGTGA
TAAGACAAGCAAAGCAGCTGGAAAAAGAATAGTCTTTATGGAGATGAAACCTGGGTTAAATTA
TTCCCAAAGCATTGTGGAATATGATGGAACACCTCATTTCGTGTAGATTACACAGAGGT
GGATAATAATGTCACGAGGCATTGGATAAAGTATTAAAAGAGGAGATTGGGACATATTAATCC
TCCACTACCTGGGCTGGACCACATTGCCACATTTCAGGGCCAACAGCCCCCTGATTGGGAG
AAGCTGAGCGAGATGGACAGCGTGTGATGAAGATCCACACCTCACTGCAGTCGAAGGGAGAGA
GACGCCTTACCCAAATTGCTGGTTCTTGTGGTGAACATGGCATGTCTGAAACAGGAAGTCACG
GGGCCTCCTCACCGAGGGAGGTGAATACACCTCTGATTTAATCAGTTCTGCCTTGAAAGGAAA
CCCGGTGATATCCGACATCCAAAGCACGTCCAAATGACGGATGTGGCTGCACACTGGCATAGC
ACTTGGCTTACCGATTCCAAAAGACAGTGTAGGGAGCCTCTATTCCAGTTGTGGAAGGAAGAC
CAATGAGAGAGCAGTTGAGATTTCACATTGAATACAGTCAGCTTAGTAAACTGTTGCAAGAG
AATGTGCCGTATATGAAAAGATCCTGGTTGAGCAGTTAAAATGTCAAGAAAGATTGCATGG
GAACGGATCAGACTGTACTGGAGGAAAGCATTAGAAGTCAGCTGAGCTGTCCCTGAGTGCACAAGT
TTCTCAGGCAGTACCTGGATGCTCTGAAAGACGCTGAGCTGTCCCTGAGTGCACAAGT
TTCTCACCCCTGCTCCTGCTCAGCGTCCACAGGCACTGCACAGAAAGGCTGAGCTGGAAAGTCCA
CTGTCATCTCCTGGTTCTGCTCTTATTGGTGATCCTGTTCAG
CATTGTGTGCACCTCAGCTGAAAGTTGCTACTTCTGTGGCTCTCGTGGCTGGCGAGGCT
GCCTTCGTTACCAGACTCTGGTTGAACACCTGGTGTGCCAAGTGCTGGCAGTGCCTGGAC
AGGGGGCCTCAGGGAAAGGACGTGGAGCAGCCTTATCCCAGGCCTGGGTGTCCCACACAGGTG
TTCACATCTGTGCTGTCAGGTCAGCTCAGTTCAG
CAAGGTGATTGTAAAGAGCTGGCGTCACAGAGACAGCCCCAGT
TCGGACAGCCTCCCAGCAGAGGTGTGGAGCTGCAGCTGGGAAGAAGAGACAATCGGCCTGGA
CACTCAGGAGGGTCAAAGGAGACTTGGTCGCACCACTCATCCTGCCACCCCCAGAATGCATCCT
GCCTCATCAGGTCCAGATTCTTCAAGGCGGACGTTCTGTTGAATTCTAGTCCTGGCC
TCGGACACCTCATTGTTAGCTGGGAGTGGTGGTGAAGCAGTGAAGAAGAGGCGGATGGTCAC
ACTCAGATCCACAGAGCCCAGGATCAAGGGACCCACTGCAGTGGCAGCAGGACTGTTGGCCCC
ACCCCAACCCCTGCACAGCCCTCATCCCTCTGGCTTGAGCCGTCAGAGGCCTGAG
CTGACCGAGACACTCACAGCTTGTCATCAGGCACGGTCAT
CCACGCTTGCACCTCGGGCCATCTGGGCTCATGCTCTCTCTGCTATTGAATTAGTACCTAG
CTGCACACAGTATGTAGTTACAAAAGAATAACGGCAATAATTGAGAAAAAAA

FIGURE 140

MRLGSGTFATCCVAIEVLGIAVFLRGFFPAPVRSSARAEGAEPPAPEPSAGASSNWTTLPPPLF
SKVVIVLIDALRDDFVFGSKGVKFMPYTTYLVEKGASHSFVAEAKPPTVTMPRIKALMTGSLPGF
VDVIRNLNSPALLEDSVIRQAKAAGKRIVFYGDETWVKLFPKHFVEYDGTSFFVSDYTEVDNNV
TRHLDKVLKRGDWDLILYLHIGLHDHIGHISGPNSPLIGQKLSEMDSVLMKIHTSLQSKERETPLP
NLLVLCGDHGMSETGSHGASSTEEVNTPLILISSAERKPGDIRHPKHQ

Important features of the protein:

Signal peptide:

amino acids 1-34

Transmembrane domain:

amino acids 58-76

N-glycosylation sites.

amino acids 56-60, 194-198

N-myristoylation sites.

amino acids 6-12, 52-58, 100-106, 125-131, 233-239, 270-276,
275-281, 278-284

Amidation site.

amino acids 154-158

Cell attachment sequence.

amino acids 205-208

FIGURE 141

GGCACGAGGCAAGCCTTCCAGGTTATCGTGACGCACCTGAAAGTCTGAGAGCTACTGCCCTACA
GAAAGTTACTAGTGCCTAAAGCTGGCGCTGGCACTGATGTTACTGCTGCTGGAGTACAACT
TCCCTATAGAAAACAACTGCCAGCACCTTAAGACCACTACACACCTCAGAGTGAAGAACTTAAAC
CCGAAGAAATTCAGCATTCATGACCAGGATCACAAAGTACTGGTCTGGACTCTGGGAATCTCAT
AGCAGTTCCAGATAAAAACTACATACGCCCAGAGATCTTCTTGCAATTAGCCTCATCCTGAGCT
CAGCCTCTGCGGAGAAAGGAAGTCCGATTCTCCTGGGGTCTCTAAAGGGAGTTTGTCTAC
TGTGACAAGGATAAAGGACAAAGTCATCCATCCCTCAGCTGAAGAAGGAGAAACTGATGAAGCT
GGCTGCCAAAAGGAATCAGCACGCCGCCCTCATCTTATAGGGCTCAGGTGGCCTGG
ACATGCTGGAGTCGGCGCTACCCCGATGGTCATCTGCACCTCCTGCAATTGTAATGAGCCT
GTTGGGTGACAGATAAATTGAGAACAGGAAACACATTGAATTTCATTTCAACCAGTTGCAA
AGCTGAAATGAGCCCCAGTGAGGTCAGCGATTAGGAAACTGCCATTGAACGCCCTCGCTA
ATTGAACTAATTGTATAAAACACCAAACCTGCTCACT

FIGURE 142

MLLLLLEYNFPPIENNQHLKTTHTFRVKNLNPKKFSIHDQDHKVLVLDGNLIAVPDKNYIRPEI
FFALASSLSSASAEGSPILLGVSKGEFCLYCDKDKGQSHPSLQLKKEKLMKLAQKESARRPFI
FYRAQVGWSNMLESAAHPGWFIGTSCNCNEPVGVTDKFENRKHIEFSFQPVCKAEMSPSEVSD

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 33-36

N-myristoylation site.

amino acids 50-55, 87-92

Interleukin-1

amino acids 37-182

FIGURE 143

CTAGAGAGTATAGGGCAGAAGGATGGCAGATGAGTGACTCCACATCCAGAGCTGCCTCCCTTAA
TCCAGGATCCTGTCCTCCTGCTGTAGGAGTGCCTGCCAGTGTGGGGTGAGACAAGTTG
TCCCACAGGGCTGTCTGAGCAGATAAGATTAAGGGCTGGGTCTGCTCAATTAACTCCTGTGG
CACGGGGCTGGGAAGAGCAAAGTCAGCGGTGCCTACAGTCAGCACCATGCTGGCCTGCCGTGG
AAGGGAGGTCTGTCCTGGCGCTGCTGCTGCTCTCTTAGGCTCCAGATCCTGCTGATCTATGC
CTGGCATTCCACGAGCAAAGGGACTGTGATGAACACAATGTCATGGCTCGTTACCTCCCTGCCA
CAGTGGAGTTGCTGCCACACATTCAACCAACAGAGCAAGGACTACTATGCCCTACAGACTGGGG
CACATCTGAATTCCCTGGAAGGAGCAGGTGGAGTCCAAGACTGTATTCTCAATGGAGCTACTGCT
GGGGAGAACTAGGTGTGGAAATTGAAGACGACATTGACAACGCCATTCCAAGAAAGCACAG
AGCTGAACAAACTTCACCTGCTTCTCACCACAGCACCAGGCCCTGGATGACTCAGTCAGC
CTCCTGAACAAAGACCTGCTGGAGGGATTCCACTTGAGTGAAACCCACTCACAGGCTGTCCATGT
GCTGCTCCACATTCCGTGGACATCAGCACTACTCTCCTGAGGACTCTCAGTGGCTGAGCAGCT
TTGGACTTGTGTTATCCTATTTGCATGTGTTGAGATCTCAGATCAGTGTGTTAGAAAATCC
ACACATCTTGAGCCTAATCATGTAGTAGATCATTAAACATCAGCATTAAAGAAAAAAAAAAA
AAA

FIGURE 144

MLGLPWKGGLSWALLLGSQILLIYAWHFHEQRDCDEHNMARYLPATVEFAVHTFNQQSKDY
YAYRLGHILNSWKEQVESKTVFSMELLGRTRCGKFEDDIDNCHFQESTELNNTFTCFFTISTRP
WMTQFSLLNKTCLEGFH

Important features of the protein:

Signal peptide:

amino acids 1-25

N-glycosylation sites.

amino acids 117-121, 139-143

N-myristoylation site.

amino acids 9-15

FIGURE 145

CTGTGCAGCTCGAGGCTCCAGAGGCACACTCCAGAGAGCCAAGGTTCTGACGCG**GATGAGGAAG**
CACCTGAGCTGGTGGCTGCCACTGTCTGCATGCTGCTCTCAGCCACCTCTGCGGTCCA
GACGAGGGGCATCAAGCACAGAATCAAGTGGAACCGGAAGGCCCTGCCAGCACTGCCAGATCA
CTGAGGCCAGGTGGCTGAGAACCGCCGGAGCCTCATCAAGCAAGGCCGCAAGCTGACATT
GACTTCGGAGCCGAGGGCAACAGGTACTACGAGGCCAACTACTGGCAGTTCCCCGATGGCATCCA
CTACAACGGCTGCTCTGAGGCTAATGTGACCAAGGAGGCATTGTCACCGGCTGCATCAATGCCA
CCCAGGCAGCGAACCAAGGGGAGTTCCAGAACAGACAAGCTCCACCAGCAGGTGCTCTGG
CGGCTGGTCCAGGAGCTCTGCTCCCTCAAGCATTGCGAGTTGGAGAGGGCGCAGGACT
TCGGGTACCATGCACCAGCCAGTGCCTCTGCCTTCTGGCTTGTGATCTGGCTCATGGT**GAAAT**
AAGCTGCCAGGAGGCTGGCAGTACAGAGCGCAGCAGCAGCAAATCTGGCAAGTGA**CCAGCT**
CTTCTCCCCAAACCCACCGCTGTTCTGAAGGTGCCAGGAGCGCGATGCACTCGCACTGCAA
TGCCGCTCCACGTATGCGCCCTGGTATGTGCCTGCGTTCTGATAGATGGGGACTGTGGCTTCT
CCGTCACTCCATTCTCAGCCCTAGCAGAGCGTCTGGCACACTAGATTAGTAGTAAATGCTTGAT
GAGAAGAACACATCAGGCACTGCGCCACCTGCTCACAGTACTTCCAACAACCTTAGAGGTAG
GTGTATTCCCGTTTACAGATAAGGAAACTGAGGCCAGAGAGCTGAAGTACTGCACCCAGC
ACCAGCTAGAAAGTGGCAGAGCCAGGATTCAACCCCTGGCTTGTCTAACCCAGGTTCTGCTCT
GTCCAATTCCAGAGCTGTCTGGTGTACTTATGTCTCACAGGGACCCACATCCAACATGTAT
CTCTAATGAAATTGTGAAAGCTCCATGTTAGAAATAATGAAAACACCTGA

FIGURE 146

MRKHLSWWLATVCMLLFSHLSAVQTRGIKHRIKWNRKALPSTAQITEAQVAENRPGAFIKQGRK
LDIDFGAEGNRYYEANYWQFPDGIHYNGCSEANVTKEAFVTGCINATQAANQGEFQKPDNKLHQO
VLWRLVQELCSLKHCFLWLERGAGLRTMHQPVLLCLLALIWLKV

Important features of the protein:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 157-171

N-glycosylation sites.

amino acids 98-102, 110-114

Tyrosine kinase phosphorylation site.

amino acids 76-83

N-myristoylation sites.

amino acids 71-77, 88-94, 93-99, 107-113, 154-160

Amidation site.

amino acids 62-66

FIGURE 147

GCCTTGGCCTCCAAAGGGCTGGGATTATAGGCGTGACCACCATGTCTGGTCCAGAGTCTCATTT
CCTGATGATTATAGACTCAAAGAAAACTCATGTTCAGAAGCTCTTCTCTGGCCTCCTCT
CTGTCTTCTTCCCTTTCTTCTTATTAAATTAGTAGCATCTACTCAGAGTCATGCAAGCTGG
AAATCTTCATTTGCTTGTCAAGTGGGTAGGTCACTGAGTCTTAGTTTTATTGGAAATT
CAACTTCAGATTCAAGGGGTACATGTGAAGGTTGTTATGAGTATATTGCATGATGCTGAGG
TTTGGGGT

FIGURE 148

MFRSSLLFWPPLCLLSLFLLLILISSIYSESCKLEIFHFACQWGRSLSLSFYFLKFQLSDSGGTCE
GLFYEYIA

Important features of the protein:

Signal peptide:

amino acids 1-25

N-myristoylation site.

amino acids 62-68

FIGURE 149

GTCTCCGCGTCACAGGAACCTCAGCACCCACAGGGCGGACAGCGCTCCCTCTACCTGGAGACTTGAC
TCCCGCGGCCAACCCCTGCTTATCCCTGACCGTCGAGTGTCAAGAGATCCTGCAGCCGCCAGTCC
CGGCCCCCTCTCCC GCCCCACACCCACCCCTGGCTCTTCTGTTTACTCCTCCTTTCAATTCTATA
ACAAAAGCTACAGCTCCAGGAGCCCAGGCCGGCTGTGACCCAAGCCGAGCGTGGAAAGAATGGGTT
CCTCGGGACCGGCACTGGATTCTGGTGTAGTGCCTCCGATTCAAGCTTCCCAAACCTGGAGGAA
GCCAAGACAAATCTCTACATAATAGAGAATTAAGTGCAGAAAGACCTTGAATGAACAGATTGCTGAA
GCAGAAGAAGACAAGATTAACATATCCTCCAGAAAACAAGCCAGGTCAAGAGCAACTATTCTT
TGTTGATAACTTGAACCTGCTAAAGGCAATAACAGAAAAGGAAAAATTGAGAAAGAAAGACAATCTA
TAAGAAGCTCCCCACTTGATAATAAGTTGAATGTGGAGATGTTGATTCAACCAAGAATCGAAAATG
ATCGATGATTATGACTCTACTAAGAGTGGATTGGATCATAAATTCAAGATGATCCAGATGGTCTTCA
TCAACTAGACGGACTCCTTAACCGCTGAAGACATTGTCCATAAAATCGCTGCCAGGATTATGAAG
AAAATGACAGAGCCGTGGACAAGATTGTTCTAAACTACTTAATCTCGGCCTTATCACAGAAAGC
CAAGCACATACACTGGAAGATGAAGTAGCAGAGGTTACAAAATTAATCTCAAAGGAAGCCAACAA
TTATGAGGAGGATCCAAATAAGCCACAAGCTGGACTGAGAATCAGGCTGGAAAATACCAGAGAAAG
TGACTCCAATGGCAGCAATTCAAGATGGCTTGCTAAGGGAGAAAACGATGAAACAGTATCTAACACA
TTAACCTTGACAAATGGCTGGAAAGGAGAACTAAAACCTACAGTGAAGACAACCTTGAGGAACCTCA
ATATTCCCAAATTCTATGCGCTACTGAAAAGTATTGATTCAAGAAAAGAAGCAAAAGAGAAAGAAA
CACTGATTACTATCATGAAAACACTGATTGACTTTGTGAAGATGATGGTAAATGCTCTCAGACCAAAACAAGCT
CCAGAAGAAGGTGTTCTACCTTGAAAACCTGGATGAAATGATTGCTCTCAGACCAAAACAAGCT
AGAAAAAAATGCTACTGACAATATAAGCAAGCTTCCCAGCACCATCAGAGAAGAGTCATGAAGAAA
CAGACAGTACCAAGGAAGAAGCAGCTAACGATGGAAAAGGAATATGGAAGCTTGAAGGATTCCACAAAA
GATGATAACTCCAACCCAGGAGGAAAGACAGATGAACCCAAAGGAAAACAGAAGCCTATTGGAAGC
CATCAGAAAAAAATTGAATGGTGAAGAACATGACAAAAAGGAAATAAGAAGATTATGACCTT
CAAAGATGAGAGACTTCATCAATAACAGCTGATGCTTATGTGGAGAAAGGCATCCTTGACAAGGAA
GAAGCCGAGGCCATCAAGCGATTAGCAGCCTGAAAAAATGGCAAAAGATCCAGGAGTCTTCAA
CTGTTCAAGAAAACATAATATAGCTAAAACACTTCTAATTCTGTGATTAAAATTTTGACCCAAGG
GTTATTAGAAAGTGCTGAATTACAGTAGTTAACCTTTACAAGTGGTAAAACATAGCTTCTTCCC
GTAAAAACTATCTGAAAGTAAAGTTGTATGTAAGCTGAAAAA
AAAAAAAAAAAAA

FIGURE 150

MGFLGTGTWILVLVLPIQAFPKPGGSQDKSLHNRELSAERPLNEQIAEAEEDKIKKTYPPENKPG
QSNYSFVDNLNLKAITEKEKIEKERQSIRSSPLDNKLNVEDVDSTKNRKLIDDYDSTKSGLDHK
FQDDPDGLHQLDGTPPLAEDIVHKIAARIYEENDRAVFDKIVSKLLNLGLITESQAHTLEDEVAE
VLQKLISKEANNYEEDPNKPTSWTENQAGKIPEKTPMAAIQDGLAKGENDETWSNTLTNGLE
RRTKTYSEDNFEELQYFPNFYALLKSIDSEKEAKEKETLITIMKTLIDFVKMMVKYGTISPEEGV
SYLENLDEMIALQTKNKNATDNISKLFPAPSEKSHEETDSTKEEAKMEKEYGSLKDSTKDD
NSNPGGKTDEPKGKTEAYLEAIRKNIEWLKKHDKGKGNKEDYDLSKMRDFINKQADAYVEKGILD
EEAEAIKRIYSSL

N-glycosylation sites:

amino acids 68-71, 346-349, 350-353

Casein kinase II phosphorylation site:

amino acids 70-73, 82-85, 97-100, 125-128, 147-150, 188-191, 217-
220, 265-268, 289-292, 305-308, 320-323, 326-329, 362-365, 368-
341, 369-372, 382-385, 386-389, 387-390

N-myristoylation sites:

amino acids 143-148, 239-244

FIGURE 151

CGGCTCGAGGCTCCGCCAGGAGAAAGAACATTCTGAGGGAGTCTACACCCGTGGAGCTCAA
GATGGTCCTGAGTGGGGCGCTGTGCTTCCGAATGAAGGACTCGGCATTGAAGGTGCTTATCTGC
ATAATAACCAGCTCTAGCTGGAGGGCTGCATGCAGGGAAAGTCATTAAAGGTGAAGAGATCAGC
GTGGTCCCCAATCGGTGGCTGGATGCCAGCCTGTCATCCTGGGTGCTCAGGGTGGAAAG
CCAGTGCCTGTCATGTGGGTGGGCAGGAGCCACTCTAACACTAGAGCCAGTGAACATCATGG
AGCTCTATCTTGGTCCAAGGAATCCAAGAGCTCACCTCTACCCGGGACATGGGGCTCACC
TCCAGCTCGAGTCGGCTGCCAACCCGGCTGGTCTGTCACGGTGCCTGAAGCCGATCAGCC
TGTCAAGACTCACCCAGCTTCCGAGAATGGTGGCTGGAATGCCCATCACAGACTCTACTTCC
AGCAGTGTGACTAGGGCAACGTGCCCCCAGAACCTCCCTGGGAGGCCAGCTCGGGTGGGGT
GAGTGGAGGAGACCCATGGCGACAATCACTCTCTGCTCTCAGGACCCCCACGTCTGACTTAG
TGGGCACCTGACCACTTGTCTTGTGGTCCCAGTTGGATAAATTCTGAGATTGGAGCTCAGT
CCACGGTCCTCCCCACTGGATGGTGCTACTGCTGTGAAACCTTGTAACCAACATGTGGGTAAA
CTGGGAATAACATGAAAAGATTCTGTGGGGTGGGTGGGGAGTGGTGGGAATCATTCTGCT
TAATGGTAACTGACAAGTGTACCCCTGAGCCCCGAGGCCAACCCATCCCCAGTTGAGCCTTATA
GGGTCAAGTAGCTCTCACATGAAGTCTGTCACTCACCCTGTGCAAGGAGAGGGAGGTGGTCATA
GAGTCAGGGATCTATGGCCCTGGCCCAGCCCCACCCCCCTTCAATCCTGCCACTGTCTA
TGCTACCTTCCTATCTCCCTCATCATCTTGTGGCATGAGGAGGGAGGTGGTGTAGAA
GAAATGGCTCGAGCTCAGAAGATAAAAGATAAGTAGGGTATGCTGATCCTCTTTAAAACCAA
GATAACATCAAATCCCAGATGCTGGTCTCTATCCCATGAAAAGTGTCTGACATATTGAGA
AGACCTACTTACAAAGTGGCATATTGCAATTATTAAATTAAAGATAACCTATTATATT
TCTTATAGAAAAAGTCTGGAAGAGTTACTTCAATTGTAGCAATGTCAGGGTGGCAGTAT
AGGTGATTTCTTTAATTCTGTTAATTATCTGTATTCCTAATTCTTACAATGAAGATGA
ATTCTTGTATAAAATAAGAAAAGAAATTAAATCTGAGGTAAAGCAGAGCAGACATCATCTGA
TTGTCTCAGCCTCCACTCCCCAGAGTAAATTCAAATTGAATCGAGCTCTGCTCTGGTGG
TTGTAGTAGTGTAGCAGGAAACAGATCTCAGCAAAGCCACTGAGGAGGGCTGTGAGTTGT
GTGGCTGGAATCTGGTAAGGAACCTAAAGAACAAAATCATCTGTAATTCTTCTAGAAG
GATCACAGCCCCTGGGATTCCAAGGCATTGGATCCAGTCTCTAAGAAGGCTGCTACTGGTGA
ATTGTGCCCCCTCAAATTCACATCCTCTTGAATCTCAGTCTGTGAGTTATTGGAGATAAG
GTCTCTGAGATGTAGTTAGTTAAGACAAGGTCTGCTGGATGAAGGTAGACCTAAATTCAATAT
GACTGGTTCTGTATGAAAAGGAGAGGACACAGAGACAGAGGAGACGCCAGGGAAAGACTATGTA
AAGATGAAGGCAGAGATCGGAGTTTGAGCCACAAGCTAAGAAACACCAAGGATTGGCAACC
ATCAGAAGCTGGAAGAGGCAAAGAACATTCTCCCTAGAGGCTTCTAGAGGATAACGGCTCTG
CTGAAACCTTAATCTCAGACTCCAGCCTCTGAACGAAGAACAAATTTCGGCTTTAA
GCCACCAAGGATAATTGGTTACAGCAGCTTAGGAAACTAATACAGCTGCTAAATGATCCCTGT
CTCCTCGTCTTACATTCTGTGTGTCCTCCACAATGTACCAAAGTTGTCTTGACCAA
TAGAATATGGCAGAAGTGATGGCATGCCACTCCAAGATTAGTTATAAAAGACACTGCAGCTTC
TACTTGAGCCCTCTCTGCCACCCACGCCCAACTATCTTGTGCTCACTCGCTCTGGGG
AAGCTAGCTGCCATGCTATGAGCAGGCTATAAAGAGACTTACGTGGTAAAAATGAAGTCTCCT
GCCACAGCCACATTAGTGAACCTAGAAGCAGAGACTCTGTGAGATAATCGATGTTGTTTT
AAGTTGCTCAGTTGGTCTAATTGTATGCAGCAATAGATAAAATATGCAGAGAAAGAG

FIGURE 152

MVLSGALCFRMKDSALKVLYLHNNQLLAGGLHAGKVIKGEIISVVPNRWLDASLSPVILGVQGGS
QCLSCGVGQEPTLTLEPVNIMELYLGAKESKSFTFYRRDMGLTSSFESAAYPGWFLCTVPEADQP
VRLTQLPENGGWNAPITDFYFQQCD

N-myristoylation sites.

amino acids 29-34, 30-35, 60-65, 63-68, 73-78, 91-96, 106-111

Interleukin-1 signature.

amino acids 111-131

Interleukin-1 proteins.

amino acids 8-29, 83-120, 95-134, 64-103

FIGURE 153

CTTCAGAACAGGTTCTCCTCCCCAGTCACCAGTTGCTCGAGTTAGAATTGTCTGCAATGGCCGC
CCTGCAGAAATCTGTGAGCTTTCTTATGGGGACCCTGCCACCAGCTGCCTCCTCTCTGG
CCCTCTTGGTACAGGGAGGAGCAGCTGCCCATCAGCTCCACTGCAGGCTTGACAAGTCCAAC
TTCCAGCAGCCCTATATCACCAACCGCACCTCATGCTGGCTAAGGAGGCTAGCTTGGCTGATAA
CAACACAGACGTTCGTCTCATGGGGAGAAACTGTTCCACGGAGTCAGTATGAGTGAGCGCTGCT
ATCTGATGAAGCAGGTGCTGAACCTCACCTGAAGAAGTGCTGTTCCCTCAATCTGATAGGTT
CAGCCTTATATGCAGGAGGTGGTGCCTGCCCTGCCAGGCTCAGCAACAGGCTAACAGCACATGTCA
TATTGAAGGTGATGACCTGCATATCCAGAGGAATGTGCAAAAGCTGAAGGACACAGTGAAAAGC
TTGGAGAGAGTGGAGAGATCAAAGCAATTGGAGAACTGGATTGCTGTTATGTCTCTGAGAAAT
GCCTGCATTTGACCAGAGCAAAGCTGAAAATGAATAACTAACCCCCTTCCCTGCTAGAAATAA
CAATTAGATGCCCAAAGCGATTTTTAACCAAAAGGAAGATGGGAAGCCAAACTCCATCATG
ATGGGTGGATTCCAAATGAACCCCTGCGTTAGTTACAAAGGAAACCAATGCCACTTTGTTATA
AGACCAGAAGGTAGACTTCTAACGCATAGATATTATTGATAACATTCAATTGTAACGGTGTTC
TATACACAGAAAACAATTATTAAATAATTGTCTTTCCATAATCAGTACTTATATTATAAA
TCCTTTAGGGAAAAACCCCTAAATAGCTTCATGTTCCATAATCAGTACTTATATTATAAA
TGTATTATTATTATAAGACTGCATTTATTATCATTATTAAATATGGATTATTAT
AGAAACATCATTGATATTGCTACTTGAGTGTAAAGGCTAATTGATATTATGACAATAATTAT
AGAGCTATAACATGTTATTGACCTCAATAAACACTGGATATCCC

FIGURE 154

MAALQKSVSSFLMGTIATSCLLLALLVQGGAAAPISSHCRLDKSNFQQPYITNRTFMLAKEASL
ADNNNTDVRLIGEKLFHGVSMSERCYLMQVLNFTLEEVLFQSDRFQPYMQEVVPFLARLSNRLS
TCHIEGDDLHIQRNVQKLKDTVKKLGESGEIKAIGELDLLFMSLRNACI

Important features of the protein:

Signal peptide:

amino acids 1-33

N-glycosylation sites.

amino acids 54-58, 68-72, 97-101

N-myristoylation sites.

amino acids 14-20, 82-88

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 10-21

FIGURE 155

GGCTTGCTGAAAATAAAATCAGGACTCTAACCTGCTCCAGTCAGCCTGCTCCACGAGGCCTGT
CAGTCAGTCCCCGACTTGTGACTGAGTGTGCAGTGCCCAGCATGTACCAGGTAGTGCAGAGGGC
TGCCTGAGGGCTGTGCTGAGAGGGAGAGGAGCAGAGATGCTGCTGAGGGTGGAGGGAGGCCAAGC
TGCCAGGTTGGGGCTGGGGCCAAGTGGAGTGAGAAACTGGATCCCAGGGGGAGGGTGCAGAT
GAGGGAGCGACCCAGATTAGGTGAGGACAGTTCTCTCATTAGCCTTCTACAGGTGGTTGCAT
TCTTGCAATGGTCATGGGAACCCACACCTACAGCCACTGGCCCAGCTGCTGCCAGCAAAGGG
CAGGACACCTCTGAGGAGCTGCTGAGGTGGAGCACTGTGCCTGTGCCCTCCCTAGAGCCTGCTAG
GCCCAACC GCCACCCAGAGTCCTGTAGGCCAGTGAAGATGGACCCCTAACAGCAGGGCCATCT
CCCCCTGGAGATATGAGTTGGACAGAGACTTGAACCGGCTCCCCCAGGACCTGTACCAACGCCGT
TGCCTGTGCCCGCACTGCGTCAGCCTACAGACAGGCTCCACATGGACCCCCGGGCAACTCGGA
GCTGCTCTACCACAAACCAGACTGTCTTCTACAGGCCATGCCATGGCGAGAAGGGCACCCACA
AGGGCTACTGCCTGGAGCGCAGGCTGTACCGTGTTCCTAGCTTGTGTGTGCCGGCCCGT
GTGATGGGCTAGCCGGACCTGCTGGAGGCTGGTCCCTTTGGAAACCTGGAGGCCAGGTGTACA
ACCACTGCCATGAAGGGCCAGGATGCCAGATGCTGGCCCTGTGAAGTGCTGTGGAGCAG
CAGGATCCGGGACAGGATGGGGGCTTGGGAAAACCTGCACTTCTGCACATTTGAAAAGAG
CAGCTGCTGCTTAGGCCGCCGAAGCTGGTGTCTGTCACTTCTCAGGAAAGGTTCAA
GTTCTGCCATTTCTGGAGGCCACCACCTGTCTCTTCTCTTCCATCCCTGCTACCTG
GCCAGCACAGGCACCTTAGATATTCCCCCTGCTGGAGAAGAAAGGCCCTGGTTTATT
TGTTTGTACTCATCACTCAGTGAGCATCTACTTGGTGCATTCTAGTGTAGTTACTAGTCTT
TTGACATGGATGATTCTGAGGAGGAAGCTGTTATTGAATGTATAGAGATTATCAAATAATAT
CTTATTTAAAAATGAAAAAA

FIGURE 156

MRERPRLGEDSSLISLFLQVVAFLAMVMGHTYSHWPSCCPGQDTSEELLRWSTVPVPPLEPA
RPNRHPESCRASEDGPLNSRAISPWRYELDRDLNRLPQDLYHARCLCPHCVSLQTGSHMDPRGNS
ELLYHNQTVFYRRPCHGEKGTHKGYCLERRLYRVSLACVCVRPRVMG

Important features of the protein:

Signal peptide:

amino acids 1-32

N-glycosylation site.

amino acids 136-140

Tyrosine kinase phosphorylation site.

amino acids 127-135

N-myristoylation sites.

amino acids 44-50, 150-156

FIGURE 157

CCGGCGATGTCGCTCGTGCTAAGCCTGGCCGCGCTGTGCAGGAGCGCCGTACCCCGAGAGCC
GACCGTTCAATGTGGCTCTGAAACTGGGCCATCTCCAGAGTGGATGCTACAACATGATCTAATCC
CCGGAGACTTGAGGGACCTCCGAGTAGAACCTGTTACAACTAGTGTGCAACAGGGACTATTCA
ATTTGATGAATGTAAGCTGGGTACTCCGGGCAGATGCCAGCATCCGCTTGTGAAGGCCACCAA
GATTGTGTGACGGGAAAAGCAACTTCCAGTCCTACAGCTGTGAGGTGCAATTACACAGAGG
CCTTCCAGACTCAGACCAGACCCCTCTGGTGGTAAATGGACATTTCTACATCGGCTCCCTGTA
GAGCTGAACACAGTCTATTCTATTGGGGCCATAATATTCTAATGCAAATATGAATGAAGATGG
CCCTTCCATGTCTGTGAATTTCACCTCACAGGCTGCCTAGACCACATAATGAAATATAAAAAAA
AGTGTGTCAAGGCCGGAAGCCTGTGGATCCGAACATCACTGCTTGTAAAGAAGAATGAGGAGACA
GTAGAAGTGAACCTCACACCCTCCCCTGGAAACAGATACTGGCTCTTATCCAACACAGCAC
TATCATGGGTTTCTCAGGTGTTGAGCCACACCAGAAGAAACAAACGCGAGCTTCAGTGGTGA
TTCCAGTGACTGGGATAGTGAAGGTGCTACGGTGCAGCTGACTCCATATTCTACTTGTGGC
AGCGACTGCATCCGACATAAAGGAACAGTTGTGCTCTGCCAACAAACAGGGTCCCTTCCCT
GGATAACAACAAAAGCAAGCCGGAGGCTGGCTGCCTCTCCTCCTGCTGTCTGCTGGTGGCCA
CATGGGTGCTGGTGGCAGGGATCTATCTAATGTGGAGGCACGAAAGGATCAAGAAGACTCCCTT
TCTACCACACACTACTGCCCTTCAAGGTTCTGTGGTTACCCATCTGAAATATGTTCCA
TCACACAATTGTTACTTCACTGAATTCTCAAAACCATTGCAGAAGTGAAGTCATCCTGAAA
AGTGGCAGAAAAGAAAATAGCAGAGATGGGTCCAGTGCAGTGGCTGCCACTCAAAGAAGGCA
GCAGACAAAGTCGTTCCCTTCAATGACGTCAACAGTGTGCGATGGTACCTGTGGCAA
GAGCGAGGGCAGTCCCAGTGAGAACTCTCAAGACCTCTCCCCCTGCCTTAACCTTCTGCA
GTGATCTAAGAACGCCAGATTCTGACAAATACGTGGTGGTCTACTTAGAGAGATTGATACA
AAAGACGATTACAATGCTCTCAGTGTCTGCCCAAGTACCCACCTCATGAAGGATGCCACTGCTT
CTGTGCAGAACTTCTCCATGTCAAGCAGCAGGTGTCAGCAGGAAAAGATCACAAGCCTGCCACG
ATGGCTGCTGCTCCTTGTAG

FIGURE 158

MSLVLLSLAALCRSAVPREPTVQCGSETGPSPEWMLQHDLIPGDLRDLRVEPVTSVATGDYSILMNWSWV
LRADASIRLLKATKICVTGKSNFQSYSCVRCNYTEAFQTQTRPSGGKWTFSYIGFPVELNTVYFIGAHNIP
NANMNEGPSMSVNFTSPGCLDHIMKYKKKCVKAGSLWDPNITACKNEETVEVNFTTPLGNRYMALIQH
STIIGFSQVFEPHQKKQTRASVVIPTGDSSEGATVQLTPYFPTCGSDCIRHKGTVVLCPQTGVPFPLDNNK
SKPGGWLPLLLLSLLVATWVLVAGIYLMWRHERIKKTSFSTTLLPPIKVLVVYPSEICFHHTICYFTEFL
QNHCRCSEVILEKWQKKKIAEMGPVQWLATQKKAADKVVFLSNDVNSVCDGTCGKSEGSPSENSQDLFPLA
FNLFCSDLRSQIHLHKYVVVYFREIDTKDDYNALSVCPKYHLMKDATAFCAELLHVQQVSAGKRSQACHD
GCCSL

Important features of the protein:

Signal peptide:

amino acids 1-14

Transmembrane domain:

amino acids 290-309

N-glycosylation sites.

amino acids 67 - 71, 103 - 107, 156 - 160, 183 - 187, 197 - 201 and 283
- 287

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 228 - 232 and 319 - 323

Casein kinase II phosphorylation sites.

amino acids 178 - 182, 402 - 406, 414 - 418 and 453 - 457

N-myristoylation site.

amino acids 116-122

Amidation site.

amino acids 488-452

FIGURE 159

AGCCACCAGCGAACATGACAGTGAAGACCCTGCATGGCCCAGCCATGGTCAAGTACTTGCTGCT
GTCGATATTGGGGCTGCCTTCTGAGTGAGGCGGCAGCTCGAAAATCCCCAAAGTAGGACATA
CTTTTTCCAAAAGCCTGAGAGTTGCCCGCCTGCCAGGAGGTAGTATGAAGCTTGACATTGGC
ATCATCAATGAAAACCAGCGCGTTCCATGTCACGTAACATCGAGAGCCGCTCACCTCCCCCTG
GAATTACACTGTCACTTGGGACCCCAACCGTACCCCTCGGAAGTTGTACAGGCCAGTGTAGGA
ACTTGGGCTGCATCAATGCTCAAGGAAAGGAAGACATCTCCATGAATTCCGTTCCATCCAGCAA
GAGACCCCTGGTCGTCCGGAGGAAGCACCAAGGCTGCTCTGTTCTTCCAGTTGGAGAAGGTGCT
GGTGACTGTTGGCTGCACCTGCGTCACCCCTGTCATCCACCATGTGCAGTAAGAGGTGCATATCC
ACTCAGCTGAAGAAG

FIGURE 160

MTVKTLHGPAMVKYLLSILGLAFLSEAAARKIPKGHTFFQKPESCPVPGGSMKLDIGIINEN
QRVSMSRNIESRSTSPWNYTVTWDPNRYPSEVVQAQCRNLGCINAQGKEDISMNSVPIQQETLVV
RRKHQGCSVSFQLEKVLVTVGCTCVTPVIHHVQ

Signal sequence:

amino acids 1-30

N-glycosylation site.

amino acids 83-87

N-myristoylation sites.

amino acids 106-111, 136-141

FIGURE 161

ACACTGGCAAACAAAACGAAAGCACTCCGTGCTGGAAGTAGGAGGAGTCAGGACTCCCAGG
ACAGAGAGTGACAAACTACCCAGCACAGCCCCCTCCGCCCTCTGGAGGCTGAAGAGGGATTC
CAGCCCCCTGCCACCCACAGACACAGGGCTGACTGGGTGTCGCCCTGGTCTGGGGGGGGCAGCAC
AGGGCCTCAGGCCTGGGTGCCACCTGGCACCTAGAAGATGCCTGTGCCCTGGTCTTGCTGTCCT
TGGCACTGGGCCAAGCCCAGTGGTCTTCTCTGGAGAGGCTGTGGGGCTCAGGACGCTACC
CACTGCTCTCCGGGCCTCTCCTGCCCTCTGGGACAGTGACATACTCTGCCCTGCCCTGGGACAT
CGTGCCTGCCCTGGGGCCCTACGCACCTGCAGACAGAGCTGGTCTGAGGTGCC
AGAAGGAGACCAGACTGTGACCTCTGTCTGCGTGTGGCTGTCCACTTGGCGTGCATGGCACTGG
GAAGAGCCTGAAGATGAGGAAAAGTTGGAGGAGCAGCTGACTCAGGGTGGAGGAGCCTAGGAA
TGCCCTCTCCAGGCCAAGTCGTGCTCTCCAGGCCCTACCCACTGCCGCTGCGTCCTGC
TGGAGGTGCAAGTGCTGCCCTGTGCAAGTTGGTCAGTCTGTGGCTCTGTGGTATATGAC
TGCTTCGAGGCTGCCCTAGGGAGTGAGGTACGAATCTGGCCTATACTCAGCCCAGGTACGAGAA
GGAACTCACCCACACACAGCAGCTGCCCTGCCCTGGCTAACGTGTACAGCAGATGGTACA
ACGTGCATCTGGTCTGAATGTCTTGAGGAGCAGCACTCAGGCCCTCCCTGTACTGGAATCAG
GTCCAGGGCCCCAAAACCCGGTGGCACAAAACCTGACTGGACCGCAGATCATTACCTTGAA
CCACACAGACCTGGTCCCTGCCCTGTATTCAAGGTGTGGCCTCTGGAACCTGACTCCGTTAGGA
CGAACATCTGCCCTCAGGGAGGACCCCCCGCGCACACCAGAACCTCTGGCAAGCCGCCGACTG
CGACTGCTGACCCCTGAGAGCTGGCTGGACCGACCGTGTGCCCTGGCAGAACGGCA
GTGCTGGCAGGCTCCGGTGGGACCCCTGCCAGCCACTGGTCCCACCGCTTCCGGAGAACG
TCACTGTGGACAAGGTCTCGAGTTCCATTGCTGAAAGGCCACCCTAACCTCTGTGTTCAAGGTG
AACAGCTCGAGAAGCTGCAGCTGCAGGAGTGCTGTGGCTGACTCCCTGGGACCTCTCAAAGA
CGATGTGCTACTGTGGAGACACGAGGCCCCCAGGACAACAGATCCCTCTGTGCCCTGGAACCCA
GTGGCTGTACTTCACTACCCAGCAAAGCTCCAGGGCAGCTGCCCTGGTGGAGAGTACTTACTA
CAAGACCTGCAGTCAGGCCAGTGTCTGCAGCTATGGGACGATGACTTGGGAGCGCTATGGCCTG
CCCCATGGACAAATACATCCACAAGCGCTGGGCCCTCGTGTGGCTGCCACTCTTGGCG
CTCGCCTTCCCTCATCCTCTCTCAAAAGGATCACCGAAAGGGTGGCTGAGGCTTGGAAA
CAGGACGTCCGCTCGGGGCGGCCAGGGCCGCGCAGGCTCTGCTCCTACTCAGCGATGA
CTCGGTTTCGAGCGCTGGCCCTGGTGGCGCCCTGGCGTGGCCCTGTGCCAGCTGCCGCTGCGTGG
CCGTAGACCTGTGGAGGCCGTGAACTGAGCGCGCAGGGGCCGTGGCTTGGTTACGCGCAG
CGGCCAGACCCCTGCAGGAGGGCGCGTGGTGGCTTGTCTCTCTCCGGTGGCGCT
GTGCAGCGAGTGGCTACAGGATGGGTGTCCGGGCCGGCGCACGGCCGACGACGCCCTCC
GCGCCTCGCTCAGCTCGTGTCCCAGCTCTGCAGGGCCGGCCGGCAGCTACGTGGGG
GCCTGCTTCGACAGGCTGCTCCACCCGGACGCCGTACCCGCCCTTCCGACCGTGCCCTGGTCTT
CACACTGCCCTCCAACTGCCAGACTTCCCTGGGGCCCTGCAGCAGCCTCGGCCCGCGTCCG
GGCGGCTCCAAGAGAGAGCGGAGCAAGTGTCGGGCCCTCAGCCAGCCCTGGATAGCTACTTC
CATCCCCGGGACTCCCGCCGGGACGCCGGGTGGGACCGGGCGGGACCTGGGGGGGA
CGGGACTTAAATAAAGGCAGACGCTGTTTCTAAAAAA

FIGURE 162

MPVPWFLLSLALGRSPVVLSLERLVGPQDATHCSPGLSRLWSDS DILCLPGDIVPAPGPVLAPTHLQTELV
LRCQKETDCDLCLRVAVHLAVHGHWEEPEDEEKFGGAADSGVVEPRNASLQAQVVL SFQAYPTARCVLLEV
QVPAALVQFGQSVGSVYDCFEAALGSEVRIWSYTQPRYEKELNHTQQLPALPWLNVSA GDGVHLVLNVS
EEQHFGLSLYWNQVQGPPKPRWHKNLTGPQIITLNHTDLVPCLCIQVWPLEPDSVRTNICPFREDPRAHQN
LWQAARLRLTLQSWL LDAPCSLPAEAALCWRAPGGDPCQPLVPPLSWENVTVDKLEFPLLKGHPNLCVQ
VNSSEKLQLQECLWADSLGPLKDDVLLLET RGPQDNRS LCALEPSGCTSLPSKASTRAARLGEYLLQDLQS
GQCLQLWDDDLGALWACPMDKYIHKRWALVWLACLLFAAALSLILLKKDHAKGWLRLKQDVRSGAAARG
RAALLLYSADDSGFERILVGALASALCQLPLRVAVDLWSRRELSAQGPVAWFHAQRQTLQEGGVVLLFSP
GAVALCSEWLQDG VSGPGAHGPHDAFRASLSCVLPDFLQGRAPGSYVGACFDRLLHPDAVPALFRTVPVFT
LPSQLPDFL GALQQPRAPRSGRLQERA EQVSRALQPALDSYFHPPGTPAPGRGVPGAGPDGT

Signal sequence:

amino acids 1-20

Transmembrane domain.

amino acids 453-475

N-glycosylation sites.

amino acids 118-121, 186-189, 198-201, 211-214, 238-241, 248-251,
334-337, 357-360, 391-394

Glycosaminoglycan attachment site.

amino acids 583-586

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 552-555

N-myristoylation sites.

amino acids 107-112, 152-157, 319-324, 438-443, 516-521, 612-617,
692-697, 696-701, 700-705

FIGURE 163

GGGAGGGCTCTGTGCCAGCCCCGATGAGGACGCTGCTGACCATCTGACTGTGGGATCCCTGGCT
GCTCACGCCCTGAGGACCCCTCGGATCTGCTCCAGCACGTGAAATTCCAGTCCAGCAACTTTGA
AAACATCCTGACGTGGACAGCGGGCAGAGGGCACCCAGACACGGTCTACAGCATCGAGTATA
AGACGTACGGAGAGAGGGACTGGGTGGCAAAGAAGGGCTGTCAAGCGGATCACCGGAAGTCCTGC
AACCTGACGGTGGAGACGGCAACCTCACGGAGCTCTACTATGCCAGGGTACCGCT
GTCAGTGCAGGGAGGCAGGTCAAGCCACCAAGATGACTGACAGGTTAGCTCTGCAGCACACTAC
CCTCAAGCCACCTGATGTGACCTGTATCTCAAAGTGAGATCGATTAGATGATTGTTCATCCTA
CCCCCACGCCAATCCGTGCAGGCATGCCACCGGCTAACCCCTGGAAGACATCTCCATGACCTG
TTCTACCACCTAGAGCTCCAGGTCAACCGCACCTACCAAATGCACCTTGGGAGGGAAAGCAGAGAGA
ATATGAGTTCTCGGCTGACCCCTGACACAGAGTTCTGGCACCACATGATTGCGTTCCA
CCTGGGCCAAGGAGAGTGCCTACATGTGCCAGTGAAAGACACTGCCAGACCGGACATGGACC
TACTCCTCTCCGGAGCCTCCTGTTCTCCATGGGCTTCGTGAGTACTCTGCTACCTGAG
CTACAGATATGTCACCAAGCCGCTGACCTCCAACTCCCTGAACGTCCAGCGAGTCTGACTT
TCCAGCCGCTGCGCTCATCCAGGAGCACGTCTGATCCCTGTCTTGACCTCAGCGGCCAGC
AGTCTGGCCAGCCTGTCCAGTACTCCAGATCAGGGTGTCTGGGACCCAGGGAGCCCGCAGGAGC
TCCACAGCGCATAAGCTGTCCAGGAGATCACCTACTTAGGGCAGCCAGACATCTCCATCCTCCAGC
CCTCCAACGTGCCACCTCCCCAGATCCTCTCCCACTGTCCTATGCCCAAACGCTGCCCTGAG
GTCGGGCCCATCCATGCACCTCAGGTGACCCCCGAAGCTCAATTCCATTCTACGCCCA
GGCCATCTCTAAGGTCCAGCCTCCTATGCCCTCAAGCCACTCCGGACAGCTGGCCTCC
CCTATGGGTATGCATGGAAGGTTCTGGCAAAGACTCCCCACTGGGACACTTCTAGTCCTAAA
CACCTTAGGCCTAAAGGTCAAGCTCAGAAAGAGGCCACCAGCTGGAAGCTGCATGTTAGGTGGCCT
TTCTCTGCAAGGAGGTGACCTCCTGGCTATGGAGGAATCCAAGAAGCAAATCATTGCAACCAGC
CCCTGGGATTGACAGACAGAACATCTGACCCAAATGTGCTACACAGTGGGAGGAAGGGACA
CCACAGTACCTAAAGGCCAGCTCCCCCTCCTCTCAGTCCAGATCGAGGGCCACCCATGTC
CCTCCCTTGCAACCTCCTCCGGTCCATGTTCCCCCTCGGACCAAGGTCCAAGTCCCTGGGCC
TGCTGGAGTCCCTGTGTCTGCAAGGATGAAGCCAAGAGGCCAGCCCTGAGACCTCAGACCTG
GAGCAGCCCACAGAACTGGATTCTCTTTCAGAGGCCTGGCCCTGACTGTGCAGTGGAGTCCTG
AGGGAAATGGGAAAGGCTTGGTGTCTCCCTGTCCCTACCCAGTGTCACATCCTGGCTGTCA
ATCCCATGCCCTGCCACACACTCTGCGATCTGGCCTCAGACGGGTGCCCTGAGAGAAAGC
AGAGGGAGTGGCATGCAAGGCCCTGCCATGGGTGCGCTCCTCACCGGAACAAAGCAGCATGATA
AGGACTGCAGCGGGGAGCTCTGGGAGCAGCTGTGTAGACAAGCGCGTGCCTGAGCCCTG
CAAGGCAGAAATGACAGTCAAGGAGGAATGCAGGGAAACTCCGAGGTCCAGAGCCCCACCTC
CTAACACCATTGAAAGTGTCTCAGGGAAATTGCTCTCCTGCCATTCTGCCAGTTTC
ACAATCTAGCTGACAGAGCATGAGGCCCTGCCTCTCTGTCTATTGTTCAAAGGTGGGAAGAGA
GCCTGGAAAAGAACCGCCCTGGAAAAGAACCGAGAAGGAGGCTGGCAGAACCGAACACCTGC
ACTTCTGCCAAGGCCAGGGCAGCAGGACGGCAGGACTCTAGGGAGGGTGTGGCCTGCAGCTCA
TTCCCAAGGCCACTGCCTGACGTTGCACGATTCACTTCAGCTCATTCTGTGATAGAACAAAGC
GAAATGCAGGTCCACCAGGGAGGAGACACACAAGCCTTCTGCAGGCAGGAGTTCACTGACCC
ATCCTGAGAATGGGTTGAAAGGAAGGTGAGGGCTGTGGCCCTGGACGGGTACAATAACACAC
TGTACTGATGTACAACATTGCAAGCTCTGCCTGGGTTCACTGGCCATCTGGCTCAAATTCCAGC
CTCACCACTACAAGCTGTGTACTCAAACAAATGAAATCAGTGCCAGAACCTCGGTTCTC
ATCTGTAATGTGGGATCATAAACACCTACCTATGGAGTTGTGGTGAAGATGAAATGAAGTCATG
TCTTAAAGTGTAAATAGTGCCTGGTACATGGCAGTGCCAATAACGGTAGCTATTAAAAA
AAAAAAA

FIGURE 164

MRTLLTILTVGSLAAHAPEDPSDLLQHVKFQSSNFENILTWDSGPEGTPDTVYSIEYKTYGERDW
VAKKGCQRI TRKSCNLTVETGNLTEYYARVTAVSAGGRSATKMTDRFSSLQHTTLKPPDVTCIS
KVRSIQMIVHPTPTPIRAGDGHRLTLEDIFHDLFYHLELQVNRTYQMHLGGKQREYEFFGLTPDT
EFLGTIMICVPTWAKESAPYMCRVKTLPDRTWTYSFGAFLFSMGFLVAVLCYLSYRYVTKPPAP
PNSILNVQRVLTFQPLRFIQEHVLIPVFDLSPGPSSLAQPVQYSQIRVSGPREPAGAPQRHSLSEIT
YLGQPDISILQPSNVPPPQILSPLSYAPNAAPEVGPPSYAPQVTPEAQFPFYAPQAISKVQPSSY
APQATPDSWPPSYGVCMEGSGKDSPTGTLSSPKHLRPKGQLQKEPPAGSCMLGGSLQEVTSIAM
EESQEAKSLHQPLGICTDRTSDPNVLHSGEEGTPQYLGQLPLLSSVQIEGHPMSLPLQPPSGPC
SPSDQGPSPWGLLESVCPKDEAKSPAETSDLEQPTELDSLFRGLALTQWES

Signal sequence.

amino acids 1-17

Transmembrane domain.

amino acids 233-250

N-glycosylation sites.

amino acids 80-83, 87-90, 172-175

N-myristoylation sites.

amino acids 11-16, 47-52, 102-107, 531-536, 565-570

FIGURE 165

TGGCCTACTGGAAAAAAAAAAAAAGTCACCCGGGCCCGCGGTGGCCACAACATGG
CTGCGCGCCGGGCTGCTCTCTGGCTGTCGTGCTGGGGCGCTCTGGTGGTCCCAGGCCAG
TCGGATCTCAGCCACGGACGGCGTTCTGGACCTCAAAGTGTGCGGGACGAAGAGTGCAGCAT
GTTAATGTACCGTGGAAAGCTTGAAGACTTCACGGGCCCTGATTGTCGTTGTGAATTAA
AAAAAGGTGACGATGTATATGTCTACTACAAACTGGCAGGGGATCCCTGAACCTTGGCTGGA
AGTGTGAACACAGTTGGATATTTCAAAAGATTGATCAAGGTACTTCATAAATACACGGA
AGAAGAGCTACATATTCCAGCAGATGAGACAGACTTGTCTGCTTGAAGGAGGAAGAGATGATT
TTAATAGTTATAATGTAGAAGAGCTTTAGGATCTTGGAACTGGAGGACTCTGTACCTGAAGAG
TCGAAGAAAGCTGAAGAAGTTCTCAGCACAGAGAGAAATCTCCTGAGGAGTCTCGGGGCGTGA
ACTTGACCTGTGCCTGAGCCGAGGCATTAGCAGACTGATTAGGAGATGGAGAAGGTGCTTCT
CAGAGAGCACCAGGGCTGCAGGGACAGCCCTCAGCTCAGGAGAGCCACCCACACACCAGCGGT
CCTGCGGCTAACGCTCAGGGAGTGCAGTCAGGACACTTTGAAGAAATTCTGCACGATAA
ATTGAAAGTGCCGGGAAGCGAAAGCAGAACTGGCAATAGTTCTCCTGCCTCGTGGAGCGGGAGA
AGACAGATGCTTACAAAGTCCTGAAAACAGAAATGAGTCAGAGAGGAAGTGGACAGTGCCTTATT
CATTACAGCAAAGGATTCGTTGGCATCAAAATCTAAGTTGTTTACAAAGATTGTTTTAGTA
CTAAGCTGCCTTGGCAGTTGCATTTGAGCAAACAAAATATTATTATTCCCTTAAGTA
AAAAAA

FIGURE 166

MAAPGLLFWLFVLGALWWVPGQSDLSHGRRFSDLKVC~~G~~DEEC~~S~~MLMYRGKALEDFTGPDCRFVN
FKKGDDVYVYYKLAGGSLELWAGSVEHSFGYFPKDLIKVLHKY~~T~~EEELH~~I~~PADETDFVC~~F~~EGGRD
DFNSYNVEELLGSLELED~~S~~VPEESKKAAEVSQHREKSP~~E~~ESRGRELDPVPEAFRADSE~~D~~GE~~G~~A
FSESTEGLQGQPSAQESH~~P~~H~~T~~SGPAANAQGVQSSL~~D~~TFEEILHDKLK~~V~~PGSESRTGNSSPASVER
EKT~~D~~AYKVLKTEM~~S~~QRGSGQCVIHYSKGFRWHQNL~~S~~LFYKDCF

Important features of the protein:

Signal peptide:

amino acids 1-22

N-glycosylation site.

amino acids 294-298

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 30-34

Tyrosine kinase phosphorylation site.

amino acids 67-76

N-myristoylation sites.

amino acids 205-211, 225-231, 277-283

Amidation site.

amino acids 28-32

FIGURE 167

CCAGGACCAGGGCGCACCGGCTCAGCCTCTCACTTGTCAAGAGGCCGGGAAGAGAAGCAAAGCGC
AACGGTGTGGTCCAAGCCGGGCTTCTGCTCGCCTCTAGGACATACACGGGACCCCTAACTTC
AGTCCCCAAACCGCGACCCCTCGAAGTCTTGAACTCCAGCCCCGACATCCACGCCGGCACAGG
CGCAGGCAGGCAGGTCCCAGGCGAAGGCAGTGCAGCAGGGGGCTGGCAGCTGGCTCGGGC
GGCAGGGAGTAGGGCCCGCAGGGAGGCAGGGAGGCTGCATATTCAAGAGTCGGGCTCGCCCTG
GGCAGAGGCCCGCCCTCGCTCCACGCAACACCTGCTGCCACCGCGCGCATGAGCCCGTGG
TCTCGCTGCTGGCGCCCGCTGCTCTGCCGCCACGGAGCCTCTGCCGCCCGTGGTCAGC
GCCAAAAGGTGTGTTGCTGACTCAAGCATCCCTGCTACAAAATGGCTACTTCCATGAAC
GTCCAGCCGAGTGGCTTCAGGAGGCACGCCCTGGCTGTGAGAGTGAGGGAGGAGTCCTCCTCA
GCCTGAGAATGAAGCAGAACAGAAGTTAATAGAGAGCATGTTGCAAAACCTGACAAAACCCGGG
ACAGGGATTCTGATGGTATTCTGGATAGGGCTTGGAGGAATGGAGATGGCAAACATCTGG
TGCCTGCCAGATCTACCACTGGTCTGATGGAAGCAATTCCAGTACCGAAACTGGTACACAG
ATGAACCTTCCTGCGGAAGTGAAAAGTGTGTTGATGTACCAACCAACTGCCAATCCTGGC
CTTGGGGTCCCTACCTTACCACTGGAAATGATGACAGGTGTAACATGAAGCACAATTATATTG
CAAGTATGAACCAGAGATAATCCAACAGCCCTGTTAGAAAAGCCTATCTTACAAATCAACCAG
GAGACACCCATCAGAATGTGGTTACTGAAGCAGGTATAATTCCAATCTAATTATGTTGTT
ATACCAACAATACCCCTGCTTACTGATACTGGTTGCTTGGAACCTGTGTTCCAGATGCT
GCATAAAAGTAAAGGAAGAACAAAAACTAGTCCAAACCAGTCTACACTGTGGATTCAAAGAGTA
CCAGAAAAGAAAGTGGCATGGAAGTATAACTCATGACTTGGTCCAGAATTGTAATTCT
GGATCTGTATAAGGAATGGCATCAGAACATAGCTTGGATGGCTGAAATCAAAGGATCTGC
AAGATGAACTGTAAGCTCCCCCTGAGGCAAATATTAAAGTAATTGTTATATGTCTATTATTCA
TTTAAAGAATATGCTGTGCTAATAATGGAGTGAGACATGCTTATTGCTAAAGGATGCACCCAA
ACTTCAAACCTCAAGCAAATGAAATGGACAATGCAGATAAAGTGTATCAACACGTCGGAGTA
TGTGTGTTAGAAGCAATTCTTTATTCTTACCTTCAAGTTGCTATAAGTTGTTACTAGTCAATGTA
TGTATATTGTTAGAATTACAGTGTGCAAAGTATTACCTTCAAGTGTATAATTGCTAAAGTGTGATAAAA
ATGAACTGTTCTAATATTATTGTCATCTCATTTCAATACATGCTCTTGATTAAAG
AAACTTATTACTGTTGTCACACTGAATTCACACACACAAATATAGTACCATAGAAAAAGTTGT
TTCTCGAAATAATTCTCATTTCAAGCTCTGCTTGGTCAATGTCTAGGAAATCTCTCAGA
AATAAGAAGCTATTCTTAAAGTGTGATATAAACCTCCTCAAACATTACTAGAGGCAAGGAT
TGTCTAATTCAATTGTGCAAGACATGTGCCTTATAATTATTAGCTTAAAGTAAACAGATT
TTGTAATAATGTAACATTGTTAATAGGTGCATAAACACTAATGCAGTCATTGAAACAAAAGAAG
TGACATACACAATATAAATCATATGTCTCACACGTTGCCTATATAATGAGAAGCAGCTCTGA
GGGTCTGAAATCAATGTGGCCCTCTTGCCTAAACAAAGATGGTTGTCGGGTTGG
ATTGACACTGGAGGCAGATAGTTGCAAAGTTAGTCTAAGGTTCCCTAGCTGTATTAGCCTCTG
ACTATATTAGTATAACAAAGAGGTCTGTGGTTGAGACCAGGTGAATAGTCACTATCAGTGTGGAG
ACAAGCACAGCACACAGACATTAGGAAGGAAAGGAACACTGAAATCGTGTGAAAATGGGTTGG
AACCACATCAGTGCATATTGATGAGGGTTGCTTGAGATAGAAAATGGTGGCTCCTT
CTGCTTATCTCCTAGTTCTCAATGCTTACGCCTGTTCTCAAGAGAAAGTGTAACTCT
CTGGTCTTCAATGTCCCTGTGCTCCTTTAACCAAAATAAGAGTTCTGTTCTGGGGAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 168

MSRVVSLLLGAALLCGHGAFCRRVSGQKVCADFHKHPCYKMAFHELSSRVSFQEARNACESE
GGVLLSLENEAEQKLIESMLQNLTKPGTGISDGDFWIGLWRNGDGQTSGACPDLYQWSDGSNSQ
YRNWYTDEPSCGSEKCVVMYHQPTANPGLGGPYLYQWNDDRCNMKHNYICKYEPEINPTAPVEK
PYLTNQPGDTHQNVVVTEAGIIPNLIYVVIPTIPLLLLILVAFGTCCFQMLHKSKGRTKTSPNQ
STLWISKSTRKESGMEV

Important features of the protein:

Signal peptide:

amino acids 1-21

Transmembrane domain:

amino acids 214-235

N-glycosylation sites.

amino acids 86-89, 255-258

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 266-269

N-myristoylation sites.

amino acids 27-32, 66-71, 91-96, 93-98, 102-107, 109-114, 140-145, 212-217